THE SEMANTIC LINK: ACTION & LANGUAGE

An investigation of relations between different cognitive domains in early development

Katharina Kaduk, Dipl. Psych.

This thesis is submitted for the degree of Doctor of Philosophy

Department of Psychology

Lancaster University

March 2017

Declaration

This thesis is my own work and no portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification at this or any other institute of learning.

Katharina Kaduk 31st March 2017

Abstract

The processing of language may derive, in part, from the ability to make sense of others' actions or gestures. Therefore, in order to understand the emergence of cognitive structures supporting language, an understanding is required of the basic aspects of semantic expectations related to goal-directed actions early in development. These expectations can act as a scaffold for later language development, with implications for vocabulary development and language comprehension. The relations of action and language have, however, not been fully and systematically explored - especially in terms of semantics. The aim of this thesis is to build knowledge in how language derives from understanding the content of action. The emphasis has been placed on semantics, with examinations utilising multiple approaches. We designed three targeted studies that looked at different aspects of age in order to index how we think language and action will interact at those time points. In order to shed light on to specific cognitive processes and organizational changes of brain activity in relation to semantics, we made use of a combination of neurophysiological methods, primarily event-related brain potentials (ERP's) and event-related oscillations (ERO's). In this context, semantic processing represents a specific application of a more general process namely that of identifying whenever something we perceive matches the predictions we have or not, based on context. For the capacity to distinguish perceived mismatches, one needs to be able to translate perceived information into meaningful concepts that are built on past and individual experiences. Neural activation that is related to semantics could be reflected in different areas across the brain via different mechanisms. In Chapter 1, the literature on infant language development, action understanding, and possible links between the two cognitive domains is reviewed and linked to semantics. Further, the objectives of this thesis are described. In Chapter 2, the semantic processing of actions at 9 months and how this processing ability may be linked to language proficiency at 9 and 18 months was investigated. In Chapter 3, the semantic representation of newly acquired words in 10-11-month-olds was measured. In Chapter 4, 2-year-olds modulation of motor systems was examined before and after the acquisition of new actions, verbs and sounds. The results of these experiments show that semantics is interwoven in action processing and in language. The implications of the results for understanding action processing in development and its relation to language are considered in Chapter 5.

Acknowledgements

Firstly, I would like to express my appreciation to the Marie Skłodowska-Curie Action Training Network funded by the European Commission for giving me this unique opportunity to pursue my research projects. I have met some incredible people along the way from which fantastic research opportunities arose. Specifically, I would like to thank Prof. Gustaf Gredebäck and Prof. Denis Mareshal for giving me the opportunity to jointly work on different projects along with Marta, Claudia, Janna, and Livia. It was great working with you while having so much fun and laughter at the same time!

Furthermore, this research would not have been possible without all those open-minded caregivers who entrusted me with their children. It was a pleasure meeting you all. One father of a participating child, Ben Tailor, was even involved in parts of my statistical analysis later down the line.

Numerous people from Lancaster have been very supportive and deserve to be mentioned: Eugenio (thank you for your patience in explaining WTools to me. I think I finally got it the 5th time round). Barrie & Dave (thank you with your technical support and help with my Matlab scripts. We had some lovely chats while programming). Gert (thank you especially with your help on study 1). Helen and Gavin (you have just been extremely understanding of my situation especially at the end stage of my PhD). The wider Babylab-Team (thank you for not burning down the Babylab while I was off duty during the last month of writing up.) Aíne (We were in this together. A perfect symbiosis if you ask me – thank you for your guidance and friendship and for actually reading parts of my thesis. I could have not wished for a better ACT partner in crime.) Uschi, Jo, Alison, Louah, Diana, Kirsty (Note: not in a particular order. I cannot list all the kind things you did for me in order to keep me sane. Thank you for absolutely everything!)

Vincent: I have known you now for more than a decade. Back in 2005, I came to work at the MPI, utterly unprepared, naïve, thinking I would work with primates. Instead, I discovered that my task was to code videos of infants. Infants? And this is how the story began… In all these years, you looked out for me, you challenged me, you believed in me when I could and simply would not. Thank you for your friendship first and foremost and for your guidance as my Doctor Father. I was under your protective wings for all this time and you taught me how to fly.

My parents: I have never seen anyone work so hard, and be so generous and loving. The feeling to be loved unconditionally and the reassurance I could always turn to you was the foundation upon which I had the confidence to take on this PhD.

Richard and my amazing daughter Isabelle: Good job team! We did this together. Rich when I needed you most – you were always there for me. Thank you for your support, I know it wasn't always easy. Izzy, you were my greatest motivator. The prospect to have more time with you again kept me going. You are by far the best thing that ever happened to me.

There are certainly many more people who helped me one way or another, and I am sorry if I omit to mention you here. It just makes me realize what a lucky person I am to have all these incredible people around me. This is what life is all about. Thank you all for being who you are. I could not have done this without you.

Table of Contents

List of Figures and Tables

Figure 3.2: Paradigm depicting the initial learning phase with (A) experimenter, (B) novel object, (C) the infant participant sitting on the caregiver's lap, and (D, E) video cameras.

Figure 3.3: Experimental paradigm shortly after learning and 24-hours later. The infant participant sat on the caregiver's lap (A) in front of an experimental monitor with video camera (B).

Figure 3.4: Experimental sequence depicting the 3-D occlusion scene. The sequence always started with an introductory phrase, followed by the word prime and ended with the object probe. Word-object pairs were either congruent or incongruent. All analyses were time locked to the first occurrence of the object.

Figure 3.5: Averaged ERPs for all infants ($N = 20$) from 12 channels within the left parietal region for the N400 (50, 51, 52-P3, 53, 54, 58, 59, 60, 61, 65, 66, 67) – top and bottom left and right parietal region (77, 78, 79, 84, 85, 86, 90, 91, 92-P4, 96, 97, 101) top and bottom right for day 1 (top) and 2 (bottom). Bottom right averaged channel group depicts the observed N400 for the incongruent condition only, within the time window of interest highlighted in grey.

Figure 3.6: Averaged ERPs for all infants ($N = 20$) from 26 channels collapsed across the two hemispheres with the analysed area highlighted in grey. Averaged ERP for day one is displayed on the top left and day 2 on the bottom left.

Figure 3.7: Theta power presented separately for day 1 and day 2 for the congruent (left) and incongruent (right) conditions for 9 frontal channels (4, 9, 10, 11, 15, 16, 18, 19, 22). A black frame ranging from 3-5 Hz from 400-1000 ms post stimulus presentation highlights the observed and analysed theta band. 3D Scalp Maps (superior view) were computed for each condition and day representing the mean of the time window and frequency range of interest. The analysed channels are highlighted in red.

Figure 3.8: Alpha power presented separately for day 1 (top) and day 2 (bottom) for the congruent (left) and incongruent (right) conditions averaged over posterior channels. A black frame ranging from 9-10 Hz from 700–1000 ms post stimulus presentation

highlights the observed and analysed alpha band. 3D Scalp Maps (posterior view) were computed for each condition and day representing the mean of the time window and frequency range of interest. The analysed channels are highlighted in blue.

Figure 3.9: Averaged ERP's for (N = 20) subjects over right parietal channels on day 2. The framed area depicts the analysed time window of interest displaying a more pronounced N400 for the incongruent condition (red line). A line was drawn at the two points either side of the time window in which the ERP crossed the border for each individual, condition and day. A measure of deflection was computed from the response curve by summing the vertical difference between the curve and a straight line between the two end-points of the curve at each of the 75-recorded samples.

Figure 3.10: Correlation scatter plot depicting the two variables N400 deviation value difference (Incongruent - Congruent) on the x-axis and alpha band power difference (Congruent - Incongruent) on the y-axis. Larger relative values for the N400 represent the infant's ability to detect word-object mismatches in the delayed retention phase. The same infants further displayed more negativity in the alpha range in the congruent relative to the incongruent condition.

Figure 4.1: Auditory presentation (Block 1) of the pre-training assessment. A still image was presented in the centre of the screen for the full duration of the auditory presentation of the novel verb and sound.

Figure 4.2: Block 1 at post-training assessment. Participants saw the same still image in the centre of the stimulus monitor during the presentation of the two trained, two novel and 1 familiar auditory stimuli. Within one completed block, the five auditory stimuli could be presented 25 times in total.

Figure 4.3: The top two rows represent the two novel actions, action A - performed with the index finger, Action B - performed with the index finger and the thumb and action C - the familiar action squeezing. All actions involved manipulating the medium "kinetic sand", presented from the left and right equally often. Timings for each of the video clips were the same. Each action started with a 600 ms period in which the hand hovered above the moulded sand before the action itself unfolded.

Figure 4.4: Left depicts the four blocks presented at pre-training and right the four blocks presented at post-training EEG assessment.

Figure 4.5: The top two rows depict activation for the perceived auditory sound and novel verb at pre-training assessment (A) and the three subsequent rows depict activation at post-training (B) for the now trained sound, trained verb and the control variable familiar action verb for the entire segment (-200 – 1000 ms) separate for the left and right hemisphere. Corresponding 3D Topographical Scalp Maps averaged for the analysed time window (600-800 ms) across all participants (N= 20) at 8 Hz are displayed in the middle. Area of channels included for the left and right are marked on the sensor layout to the left (C).

Figure 4.6: 3D Topographical Scalp Maps across all participants (N= 20) in 100 ms intervals for the segment length including a 200 ms baseline and post stimulus activation for 1000 ms duration post stimulus presentation at 8 Hz. The top two rows depict activation for the perceived auditory sound and novel verb at pre-training assessment and the three subsequent rows depict activation at post-training for the now trained sound, trained verb and the control variable familiar action verb.

Figure 4.7: Mu power for the left and right hemisphere separately over central electrodes to the familiar verb squeezing (grey), novel sound (blue), and novel verb yokking (red) pre-training (solid colour) and post-training (textured). Note the familiar verb derived from Block 1 at post-training. (** $p < 0.01$, * $p < 0.05$ - green highlights significance from baseline, black indicates significance between conditions).

Figure 4.8: The top three rows depict activation for familiar and two novel audio-visual materials and pre-training (A) and the two subsequent rows depict activation at post-training for the now trained language and non-language video material (B) for the entire segment (-200 – 3900 ms). Analogous of the three distinct time windows (C): TW1 (600-1400 ms - auditory cue), TW2 (1400 – 2000 ms - offset of the auditory stimulus and first hand to sand contact), and TW3 (2000-2800 ms - action execution). (D) Corresponding 3D Topographical Scalp Maps for each of the three time windows

averaged for the analysed time across all participants (N= 14) at 6-8 Hz. Area of channels included for analysis are marked on the sensor layout to the bottom left (E).

Figure 4.9: Mu power to familiar, novel (Pre-Training Assessment) and trained (Post-Training Assessment) audio-visual stimuli for each of the three analysed time windows of interest. Blue colour depicts actions paired with the non-verbal and red actions paired with the verbal referent (** $p < 0.01$, * $p < 0.05$ - green highlights significance from baseline, black indicates significance between conditions).

Table 4.1: Mean and Standard Deviation (including the total number of presented trials) and sample size for the auditory and audio-visual task separate for the pre- and post-training assessments.

Chapter 1: General Introduction

To become part of the complex adult world, human infants need to be prepared to learn from other people. The ability to understand others is crucial for the development of early human social-cognitive skills. The first part of this introductory chapter aims to highlight infants' sensitivity to social aspects of their environment as a whole and as a possible foundation for the ontogeny of action perception. However, in order to become an active member of this complex adult world, human infants, who are immersed in this environment of richly socially constructed symbolic factors, must first unravel and learn to understand the meanings behind these abstract symbols. A longstanding debate in developmental research and in developmental cognitive neuroscience is how the meaning or *semantics* of action symbols emerge and what role language plays in this process. When we refer to action symbols we refer to actions where there is semantic or meaningful content beyond mere physical movement of the body. There is content in the action that is symbolically present. For example, pointing in and of itself is meaningless as an action unless it is understood that it contains information that is symbolically encoded (Daum, Ulber, & Gredebäck, 2013). Action and language may be closely related with evidence in the field pointing towards that notion. In early development however, this relation is still unclear. Current neurophysiological measures of semantics are also not well defined in the developmental literature. Research on semantic processes generated by the N400 ERP for both cognitive domains separately, provides largely robust but also confounded evidence for the link between these two cognitive domains. A general question still to be answered is - do current measures of semantics index the same cognitive construct? Consequently, the aim of

this thesis was to utilize additional measures to reflect semantic processing in developmental cognitive neuroscience. Neurophysiological measures already provide some evidence that the mirror system relies on stored semantic knowledge about actions and events. Comparably little is known about motor involvement in relation to the ontogeny of language and action systems in early development. To date, motor resonance induced by linguistic material has not been investigated to the same extent as action observation. It can be argued that a robust way to investigate the emergence of semantic representation and involvement of motor systems during action observation and language processing is to more closely examine early development.

To date, comparably little is known about the relation between language and action and how these two conceptually different cognitive domains may or may not be interwoven. Especially in terms of semantics, a possible relation of action and language has not been fully and systematically explored. It has therefore been of primary focus for the outlined doctoral thesis to investigate how semantic information develops and how semantic representations are stored and accessed by two conceptually different cognitive domains in early development. Furthermore, within the field of developmental cognitive neuroscience, we aimed to examine a variety of semantic measures and as a consequence, we explored complementary methods. This lead us to make use of a variety of neurophysiological methods, primarily event-related brain potentials (ERP's) and event-related oscillations (ERO's). Consequently, the last part of this general introduction chapter will outline my thesis objectives.

1.1 Action and Language Development in the Social Context

1.1.1 Prenatal Development

The auditory system of a foetus becomes functional at around 25 weeks' gestation. Stimulation via speech, music and meaningful sounds from the environment are already critical for the developing neurological basis of the auditory system at this particular point in development. Thus, environmental stimulation impacts on the developing auditory cortex very early on (Graven & Brown, 2008). Prosodic features are essential for later language acquisition and evidence for prenatal learning (or formation of memories) from these sensory inputs is, for instance, reflected in the newborns' cry, which follows the prosody of their mother tongue (Mampe, Friederici, Christophe, & Wermke, 2009). A newborn's ability to later recognize the face of their own mother is theorised to have its roots in prenatal learning of their mother's voice (Sai, 2004). Zoia and colleagues (2007) showed that movement patterns of fetuses reflect intentional actions, and that advanced motor planning already occurs in utero. The observation of fetal mouth opening prior to the hand making contact with the oral cavity suggests that there is a rudimentary body schema in addition to these motor-planning capacities. Marx and Nagy (2015) further proposed the possibility that fetal responses to the external world (e.g., responses to voice or touch) serve a communicative purpose similar to that of newborns. This indicates that a foetus is born into this world already prepared to learn socially delivered information.

1.1.2 Postnatal Development

Evidence from developmental studies on infants' acting upon objects and interactions with people have shown that young infants are capable of distinguishing people from inanimate objects (e.g., Legerstee, 1992; Poulin-Dubois, Lepage, & Ferland, 1996) and

that they prefer socially salient stimuli such as human faces when contrasted with other stimuli (Morton & Johnson, 1991), with a specific preference for faces with direct gaze (e.g., Farroni et al., 2002; Senju & Johnson, 2009), voices (Blasi et al., 2011) and biological motion (e.g., Bertenthal et al., 1987; Bertenthal & Pinto, 1993). However, mere auditory or visual observation of a socially salient stimulus would only provide an infant with perceptual parameters. In order to grasp the intrinsic properties of observed agents and their actions, for instance, and to understand what those actions really mean, depends on the implementation of sensory input into one's own motor specific semantic knowledge system (e.g., Hommel, 2005; Prinz, 1990; Rizzolatti, 2005).

1.1.2a The Role of Imitation

Early imitation is thought to play a pivotal role as it provides first hand understanding of how it feels to be someone else (Meltzoff & Decety, 2003). Early work on newborns' ability to imitate facial expressions has consequently led to a re-evaluation of the stand that perception and action are independent from one another (e.g., Meltzoff & Moore, 1997). For example, 5-week-olds are more likely to imitate tongue and mouth opening gestures when performed by an adult model as compared to an object model (Legerstee, 1991). When 2-month-olds encounter objects in their visual scene, an increase in grasping movements is likely to be observed. However, as soon as they are faced with another person, infants start to respond by means of facial expressions and pre-speech noises (Richards, 1974) and react with distress, when others pose a still-face, stopping reciprocal interaction (see Mesman, van IJzendoorn, & Bakermans-Kranenburg, 2009 for a review). Meltzoff (1990) has further provided evidence for infants' ability to recognize that they themselves are being imitated by means of preferential looking and an increase in positive affect towards the imitating experimenter. This highlights the role of newborns' as interactive agents, eager to explore action effects, able to foresee

events in their environment (von Hofsten, 2004). This reciprocal imitation exchange is to some extent similar to that of a communicative exchange (e.g., Meltzoff, 1990). The above evidence consequently highlights a) that already very early on, a neonate and to some extent the prenatal foetus is aware of people and inanimate objects and b) even forms different expectations for these two classifications of environmental factors (for a review see Legerstee, 1992). For instance, only if a human evidently operated a mechanical claw, the expectancy from non-goal directed to goal directed intent shifted in 9-month olds (Hofer, Hauf, & Aschersleben, 2005). Although there is no doubt that newborn infants are indeed socially interactive individuals, newborn imitation data as well as the "like me" hypothesis is viewed as controversial within the field of developmental psychology (see, e.g., Meltzoff et al., 2017; Oostenbroek et al., 2016).

1.1.2b The Role of Joint Attention

Another important and fundamental aspect for social learning is the ability to understand and be aware that another person is looking at the same object as oneself, and that the social partner is doing so to share information (Farroni, Csibra, Simion, & Johnson, 2002). This is seen as the highest level of joint attention, also known as triadic joint attention (Tomasello, 1995). Nine month olds readily encode social cues such as eye gaze referentially, with a bias to attend to the focal point of observed gaze (Senju, Csibra, & Johnson, 2008). Joint attention has shown to be influential for object processing (Striano, Chen, Cleveland, & Bradshaw, 2006). The ability to perform overt joint attention behaviours with a social partner has also been linked to subsequent language development (e.g., Baldwin, 1993; Kuhl, Tsao, & Liu, 2003). For instance, in a longitudinal study Brooks and Meltzoff (2005) showed that gaze-following behaviour of 10-11-month-old infants was predictive of later language abilities at 18 months. Several other longitudinal studies have further confirmed a relation between the ability

to understand joint attention and later vocabulary development (e.g., Tomasello & Farrar, 1986). Infants own abilities to produce pre-verbal actions such as gestures and pointing were also linked to language development. Butterworth and Morissette (1996) reported that an earlier onset of pointing correlated with the number of different gestures and the amount of speech comprehension present at 14.4 months of age. Camaioni and colleagues (1991) further added that the amount of pointing at 12 months is predictive of speech production rates at 24 months (Camaioni, Caselli, Longobardi, & Volterra, 1991). In summary, these studies illustrate the importance of understanding others' non-verbal but highly communicative acts and the infant's own ability to make use of these tools to pave the way for language development (e.g., Iverson & Goldin-Meadow, 2005).

1.1.2c The Role of Infant-Directed Speech

Another highly social signal that indicates that the infant is being addressed within social communicative actions is infant-directed speech (IDS). This type of speech has a number of attention inducing properties (e.g., heightened intonation, speech, pitch), which were shown to promote attention to speech and word-object mapping (Ma, Golinkoff, Hourton, & Hirsch-Pasek (2011). Zangl and colleagues (2007) found increased neural activity in 6- and 13-month-olds for familiar and unfamiliar words when those words were presented in an ID, as opposed to adult-directed (AD) manner despite the absence of other social cues such as eye gaze (Zangl & Mills, 2007). Singh and colleagues (2009) tested whether IDS aided the recognition of newly acquired words in 7.5-month-olds when compared to adult-directed speech (ADS). Their findings support the notion that IDS is not just simply something infants prefer. Moreover, IDS was found to be essential for word recognition and word processing. If IDS was used to label words during training, a longer lasting implementation into memory (24-hours

later) was evident regardless of how the words were presented during test (ID or AD). However, at some point in development, young language learners will have to become more attuned to ADS. At what time in development the influence of IDS reduces in importance remains a subject for further investigation (Singh, Nestor, Parikh, & Yull, 2009).

1.1.3 Action Comprehension

Unsurprisingly, the influence of an infant's social environment drives a large proportion of an infant's development (e.g., Pascalis et al., 2002). Consequently, within the first postnatal year, an infant has learned to appreciate other peoples' actions as goal-directed with a readiness to construct action representations organized hierarchically, that is with respect to an ultimate goal. These representations allow for the prediction of the consequences of actions, including the ability to interpret and describe actions, and categorize action sequences (Sommerville & Woodward, 2005). Furthermore, recent work suggests that goal-based action representations, similar to language structures, are hierarchically organized. Infants re-enact the final goal of a modelled action, but do not necessarily reproduce the means (Gergely, Bekkering, & Kiraly, 2002). By 6 months of age, infants can infer the goal of an uncompleted action without any cues of the goal itself (Daum, Prinz, & Aschersleben, 2008) and by the time infants reach 10 months, they are capable of parsing observed sequences of continuous everyday actions along intentional boundaries in respect to the completion of an action (Baldwin, Braid, Saylor, & Clark, 2001). The use of EEG methodology has furthered understanding of the underlying neurological processes of goal-directed action perception. EEG Frequency analyses in the mu band, for instance, suggest that motor representations of perceived actions are only evident when the action is goal-directed and is directly linked to the

observers own motor expertise (e.g. Nyström, Ljunghammar, Rosander, & von Hofsten, 2011). However, it is important to note that there is not enough evidence that points to the notion that higher cognitive abilities are exclusively dependent on the re-enactment of sensory and motor representations (Caramazza, Anzellotti, Strnad, & Lingnau, 2014). For example, how the motor system discriminates between actions that are unfamiliar or those that are highly salient but lie outside an infant's motor repertoire is still unknown factors (Ní Choisdealbha & Reid, 2014).

The presence or absence of an N400 ERP component to expected and unexpected action sequences further indicates that infants' interpretations about actions are based on an early developing semantic framework (e.g., Reid et al., 2009). This knowledge about actions must go beyond mere coupling of objects and outcomes due to their consistent and simultaneous appearance (Ní Choisdealbha & Reid, 2014). These findings indicate that the origins of processing and understanding hierarchical representations are to be found in infancy and that the hierarchical nature of actions can to some extent help facilitate action predictions by means of semantic rule use.

The emergence of action processing and their driving contributors are currently not well understood. We can only speculate how or if across the lifespan action and language interact with one another and ask ourselves if these processes develop in tandem or separately from one another or whether their dependence or independence fluctuates across the lifespan (Ní Choisdealbha & Reid, 2014).

1.1.4 Language Comprehension

If we were to draw an analogy to infants' language development, we can further conclude that just as actions are not separated by clear pauses, words are equally inseparable by pauses in the speech stream. Therefore, for young language learners, it

seems apparent that verb acquisition appears to be far more complex and challenging than acquiring nouns. This is, for example, reflected in a child's first 50 words of vocabulary, with nouns playing a predominant role (Ferson et al., 1994). In order to conclude a concept from an event, infants first have to know which actions semantically belong together. As verbs are not perceptually available in the same way as nouns, this may hold a possible explanation as to why verb acquisition appears to be far more complex and challenging than acquiring nouns. When looking at adult-child interactions, for example, one can further observe that adults only use a small number of verbs. Tomasello and colleagues (1992) reported that adults use verbs more frequently to regulate a child's behaviour, anticipate events or to comment on completed actions. Typically, an infant would execute an action, which would in close temporal proximity be commented upon by a caregiver using the correspondent verb of the performed action. It has been conjectured that this simultaneous representation of both action performance and auditory information will subsequently lead to a synaptic connection between neurons in specific motor and language areas that will ultimately become stronger with time (Pulvermüller, 2005). However, Pulvermüller's claims cannot fully explain whether somatotopic activation is causal for semantic processing. It could be argued that observed activation in premotor areas is elicited by semantic processing in nonmotor specific regions independent of the motor system (Caramazza et al., 2014).

Motion verbs in particular such as *fall*, *jump*, or *dance*, are easier for children to imagine, as they can be associated with particular routines or social events (Tardiff, 1996). Action words are therefore defined by abstract semantic links between language elements and motor programs. Although motion verbs cover only a small percentage of verbs from the adult vocabulary, they are yet amongst the first verbs that a child learns (Fenson et al., 1994). Therefore, in situations where all cues for verb usage (perceptual,

linguistic, and social) meet, verb learning is being facilitated. Golinkoff and colleagues (2002) described three key steps that are needed for verb learning: a) attending towards actions and relations in the environment, b) forming categories of actions and relations without language, and c) mapping words to actions. Word mapping can of course only progress if the action has additionally been labeled. This gives the child a further cue, namely, that the action was purposeful. Such action analysis is conjectured to be central to inferring intentions. In a related sense, Gampe and Daum (2014) investigated the influence of language on action prediction in children between 12 and 24 months of age by looking at their anticipatory gaze behaviour. Their results suggest that verbs already produced by the majority of the 24-month-olds result in an enhanced action prediction, while verbs that were produced by only a minority of the children at this age did not result in any differences of the anticipation times from baseline. Because prior verb knowledge led to a faster anticipation of events, Gampe and Daum's (2014) findings are in line with the view of shared representations of the language and action domain within the first two postnatal years.

Evidence for the possible notion that sensorimotor experience for language formation becomes increasingly independent with age, as postulated by (Piaget, 1952), comes from Gampe and colleagues (2016). These authors investigated toddlers' verb learning success in two distinct conditions. Either a toddler solely observed an adult perform an action or they had the chance to produce the actions themselves. Only toddlers that imitated the adult during training of complex actions (e.g., rotating a handle on a novel box, which left the box in a distinct end state) were able to learn the corresponding action verb. Interestingly, verb learning success correlated with imitation proficiency. The authors therefore proposed that the ability to generate an action will increase the likelihood to learn a novel label. Further, only 24-30-month olds but not

36-month-olds benefited from action reproduction during the verb learning process, which led the authors to conclude that this may only be a phenomenon for early language development (Gampe, Brauer, & Daum, 2016).

1.1.5 Motor Development

Finally, the importance of infants' developing motor skills on their domain general cognitive development should not be ignored. New developmental milestones such as sitting upright automatically enhance learning within motor-specific areas. For instance, being able to judge a final end state of a grasp movement correlates with 6-month-olds own grasping performance (Daum, Prinz, & Aschersleben, 2011). Those newly mastered skills further impact on development in non-motor specific domains such as learning about objects (Soska et al., 2010). Brandone (2015), for example, used eye tracking to measure 8- to 11-month-old infants' ability to visually predict the goal of an on-going successful or failed intentional action. In addition, infants' self-locomotion and joint-attention abilities were assessed. Results support the view that joint-attention increases with infants increased ability to crawl and walk, which on the other hand may shape infants' understanding of others as being intentional agents. Object knowledge also impacts understanding on action effectors. Six-month-old infants are able to predict that a cup would go to the mouth when contrasted with predictive looks to the ear (Hunnius & Bekkering, 2010) or a spoon would be placed into a mouth (Kochukhova & Gredebäck, 2010). Motor skill development such as sitting or walking can further influence learning within the language domain (e.g., Oudgeneog-Paz et al., 2012; 2013; Walle & Campos, 2014). Libertus and Violi (2016), for instance, found sitting to be predictive of 10- and 14-month-olds subsequent receptive vocabulary size allocating particular importance to sitting due to this behaviour producing increased language

learning opportunities. In addition, developmental disorders such as language impairment have been related to motor delays (see Leonard & Hill, 2014 for a review). It is important to remember that object or action knowledge do not exclusively emerge through motor development and active experience. Human infants are remarkably fast when it comes to forming associations between events and thus observational experience remains a stable source of information about the surrounding world until such time that the infant is given the opportunity to actively perform previously observed actions themselves (Hunnius & Bekkering, 2014).

1.2 Current Measures for Action and Language Semantics

The prior overview aimed to illustrate the importance of the social context for a successful formation of language and action understanding. There is strong evidence to suggest that both domains overlap to a certain extent and have an effect on one another. To which extent these two cognitive domains are indeed dependent is currently under debate in the developmental science literature. We can ask if the relation between the two domains is unidirectional, with one domain temporarily or permanently driving the development of the other, or whether the relation between the two domains is a bidirectional one. Rüschemeyer and colleagues (2009) propose that both cognitive domains share one common ground, namely semantics, in order to function (Figure 1.1).

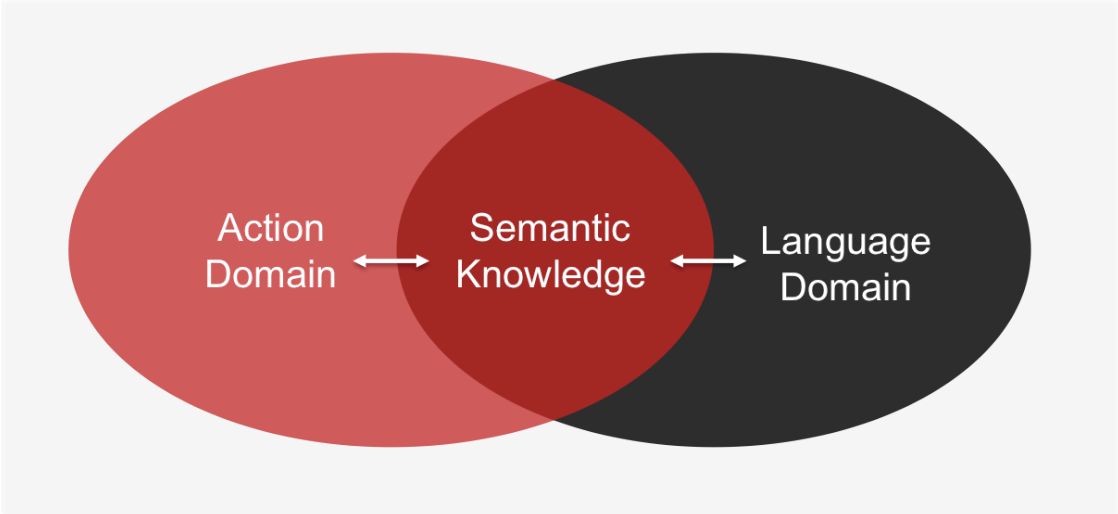

Figure 1.1: *Domain specific overlap between action and language with each domain required to access the semantic knowledge system. Modified from Rüschemeyer et al. (2009).*

Under this scenario, in order to investigate the functional properties of both domains, one must a) first look at how *meaning* is represented in each of these domains and b) make use of EEG methodology to investigate comprehension of the pre-verbal population when studying the relation between the two developing domains as they emerge during early development. The latter will also enable direct comparison with current adult findings within the mirror neuron account.

In order to become a functional member of this world, it is essential to relate incoming and outgoing abstract symbols into meaningful concepts. In the context of this thesis, when we refer to the terminology semantics, we refer to the attribution of *meaning* to linguistic and action content. In both cognitive domains, the recruitment of semantic meaning from knowledge involves constructing a mental image or representation and then comparing it with a given context (Rüschemeyer et al., 2009). The retrieved semantic knowledge thus activates expectations, which is shaped and built on past and individual experiences. These experiences, whether active (embodied actions) or passive (observed actions), will form memories. Neural activation that is

related to semantics could be reflected in different factors and functions. For instance, the activation of action related sensation/input could be reflected in different areas across the brain via different mechanisms.

In cognitive neuroscience, semantics has previously been investigated via two distinct routes. The first route will be referred to in this thesis as a more *indirect route to semantic knowledge* using the semantic expectancy paradigm. In our interpretation, this access is indirect as semantic recruitment is context dependent (i.e. context based on previous perceptual information such as a word prime). This semantic congruency pathway to semantics will be explored via event-related brain potentials, more specifically via the N400 ERP component. In both cognitive domains (action and language) this specific neural marker was found to be a sensitive indicator of the semantic relationship between the target and context with the N400 ERP morphology depicting a more negative deflection when target does not fit the expected context.

The other route infers a more *direct* access to stored semantics by means of event-related oscillations (ERO's). In comparison to the N400 literature, relatively little is known about the semantic involvement of the motor system in early development especially that of the language domain. Two distinct measures are thought to access semantics via oscillatory measures - both of which can be extracted from the alpha range of the recorded EEG. The first one represents the *upper alpha (marker for semantic congruency that is thought to reflect activation of the semantic knowledge system)* and the second *mu rhythm (evoked by action perception or execution, typically observed over the sensorimotor cortex)*. Both of these two alpha measures access semantics differently. Upper alpha, similar to the N400, is in our interpretation more context dependent and is therefore likely to be a more indirect form of measuring semantics. A more direct route to semantic memory can be realized my means of motor

resonance (mu desynchronization), as a single action stimulus can lead to the activation of motor systems. The subsequent section will provide more information for each of the neural correlates.

1.2.1 Accessing Semantics via the N400 ERP component

1.2.1a Language-N400

Derived originally from language research more than thirty years ago, one specific neural correlate was found to index processing of semantic information. For instance, when reading sentences that ended either with a congruent word (e.g., I shaved off my moustache and '*beard*') when compared to incongruent endings (e.g., I take coffee with cream and '*dog*'), adult participants in Kutas & Hillyard's (1980) study displayed a distinctly more negative deflection peaking at around 400 ms post stimulus onset for incongruent words. This component was labelled the N400 (Kutas & Hillyard, 1980; for a more in depth review, see Kutas & Federmeier, 2011). Since this discovery, the N400 event-related brain potential has been reliably replicated and linked to semantic processing for a wide array of stimulus types.

Developmental research unitizing the N400 such as work by Friedrich and Friederici (2010) found that infants with high early language word production displayed an N400 semantic priming effect in a picture-word paradigm already at 12 months. With respect to language proficiency, infants with low language production did not show this effect, even if parents rated the words to be comprehended by their children at that age. The results suggest that infants' state of behavioural language development can be reflected by the presence or absence of the N400. Thus, the N400 component is useful for studying language comprehension and it can provide insights into the way that meaning is stored in long-term memory.

1.2.1b Action-N400

Perception of action (as outlined in section 1.1) also involves semantic processes. Amoruso and colleagues (2013) suggest in their review article that both N400 measures (action and language N400) are likely to possess a common source despite being indexed by two different semantic systems (verbal in the left and image-based in the right hemisphere) (Kutas & Federmeier, 2011). Consequently, the N400 neural signature of present literature shifts toward the notion of a more boarder index for processing. The focus of the N400 as a tool for studying action processing has increasingly gained popularity in the field. In recent years, action research has made use of this neural marker to investigate how expectancy from previous experience impacts novel observed events and how they may be concluded (Reid & Striano, 2008). The N400 component was linked to sequences of action that included violation of expectation (e.g., Reid et al., 2009). Reid and colleagues (2009) found that 9-month-old infants and adults, but not 7-month-olds, produced an N400-like response when anticipating action conclusions. A speculative conclusion by the authors was that this may lead to the suggestion that language processing may derive from understanding action in early development. First, infants understand the structure of actions. Through possessing this capacity, infants likely bootstrap learning about language through the application of rules relating to the structure of action when processing language structures later in development. Reid and colleagues (2009) suggested that infants' inferences about action are based in an early developing semantic framework, which goes beyond infants' pairing of objects and outcomes simply because they always appear simultaneously (Ní Choisdealbha & Reid, 2014). In congruence with this view, van Elk and colleagues (2008) reported that planning a meaningful action as opposed to a meaningless action modulated the N400 ERP to a word, which was either related to the action's goal or not. This may indicate that semantic activation is essential for action

preparation (van Elk et al. 2008).

1.2.2 Accessing Semantics via Event-Related Oscillations

1.2.2a The Indirect route of the Upper Band of the Alpha Range

From adult literature, the work from Klimesch and colleagues (1992) is of particular interest as he argues for a distinction between lower alpha (7–10 Hz) and upper alpha (10–13 Hz) range. The lower alpha in his view represents more global neural processes mainly related to attentional demands and the latter, upper band within the alpha range is thought to be more localized and related to specific task demands in adults. In his later work, Klimesch (1999) postulates that alpha desynchronization in adults is associated with semantic long-term memory performance and theta synchronisation[1] with the encoding of new information, which is in close contact with long-term memory. Even more recently, desynchronization within the upper alpha range has further been related with the recruitment of the semantic knowledge system in adults. More specifically, the retrieval of semantically well integrated information is thought to elicits more cortical excitation when compared to less integrated information. In other words, the more incorporated information is in the semantic knowledge system, the larger the upper alpha band desynchronization (Klimesch, 2012).

In infant literature, evidence is emerging which indicates that a lower frequency band (5-7 Hz) of the alpha range reflect early social object learning mechanisms as alpha was evidently more suppressed in 4 and 9-month olds to faces with object directed eye gaze when compared to averted eye gaze (Michel et al., 2015). Hoehl and colleagues (2014) also reported desynchronization of this lower band when 9-month-

[1] Synchronisation of the theta band is associated with successful learning (e.g., Begus et al., 2015; Saby & Marshall, 2012). For instance, when there was an increase in theta power during the encoding of a new word, those words were more likely to be remembered in a subsequent retention test than those that did not display this EEG effect (Klimesch et al., 1996).

olds and adults previously engaged in eye contact prior to turning to an object presented on a computer screen when contrasted with a lack of direct eye contact. The authors postulated that eye contact leads to increased brain activation in response to an object (Hoehl et al., 2014). Whether the related brain response was related to information processing about the object or was triggered by joint attention or a mere arousal caused by the eye contact remains uncertain. To our knowledge, evidence that parallels upper alpha in adults and in early development is still unexplored.

1.2.2b The Direct route of Motor Resonance

The theoretical framework of embodied cognition proposes that language is grounded in perception and action systems (e.g., Fischer & Zwaan, 2008; Gallese, 2008; Glenberg & Gallese, 2012) with the suggestion that the simultaneous representation of both action performance and audio information consequently results in synaptic connectivity between neurons in specific motor and language areas (Pulvermüller, 2005). The role of mirror neurons [2]in the parietal and premotor cortex (e.g., Gallese, Fadiga, Fogassi, & Rizzalatti, 1996), are of specific importance as they activate during action execution as well as action observation performed by others (Fogassi et al., 2005; Rizzolatti & Craighero, 2004). Subcategories of mirror neurons could further be found in monkey premotor cortex triggered by action sounds such as ripping a piece of paper independent of the action being represented visually (Kohler, Keysers, Umiltà, Fogassi, Gallese, & Rizzolatti, 2002).

The general position present in the literature is that the equivalent to the human mirror neuron system is also reflected in mu rhythm of the recorded

[2] Mirror neurons were originally discovered in the premotor cortex of the macaque monkey and are characterized by responses produced when the animal performs an action as well as observes a similar action e.g. when the monkey is grasping an object, it would lead to the same neurons to fire as those for passively observing an experimenter's hand grasping for an object (Camarazza, Anzellotti, Strnad, & Lingnau, 2014).

electroencephalogram (EEG) (Gonzalez, Reeb-Sutherland, & Nelson, 2016). Mu rhythm desynchronization is an electrical pattern linked to motor experience typically measured over the sensorimotor cortex (Cuevas et al., 2014). Due to infants' sensitivity for understanding goals as demonstrated by numerous behavioural studies (e.g., Sommerville & Woodward, 2005) and evidence from studies looking at motor representations of perceived actions (e.g., Nyström, Ljunghammar, Rosander, & von Hofsten, 2011), it is expected that mu rhythm decreases in power (desynchronizes) when infants observe goal-directed actions or perform an action themselves when compared to mu at rest (Marshall et al., 2011).Consequently, the measure of mu rhythm is of relevance for action perception and action observation due to its responsiveness predominantly to observed actions as a function of infants' motor repertoire (e.g., Pineda, 2005; van Elk et al., 2008; Reid et al., 2011). More specifically, the suggestion from the mirror neuron account is that observed actions can only be understood if they can be mapped onto one's own motor representations (e.g., Nyström et al., 2008). Cannon and colleagues (2015) found infants with increased competencies of grasping a toy to correspond with greater mu desynchronization when observing reaching actions thus reflecting the link between perception and action systems. Nyström and colleagues (2008) found 8-month old infants to display a similar pattern of motor activation to goal-directed as opposed to non-goal-directed actions (Nyström et al., 2008). These findings are in agreement with behavioral data, which indicate that the identification of a goal from an action sequence was related to 10-month-old's own performance (Sommerville & Woodward, 2005) and that increased motor proficiency (e.g., pointing), leads to a better understanding of other's actions (Tomasello, Carpenter, & Liszkowski, 2007).

Along these lines, Saby and colleagues (2012) investigated the effect of being

imitated via mu desynchronization. Immediately after 14-month-old infants executed specific goal-directed actions (e.g., button press) they watched actions that were either in accordance with the actions that they just performed (imitating the infant) or not. Their results revealed mu desynchronization to be greater for actions that matched their own. Mayer and colleagues (2011) report greater motor activation of 3-year-olds to action observation when acting jointly in comparison to observing two others jointly performs the same action task. These findings indicate that in early toddlerhood, observing others' actions are differentially processed as a function of joint attention and are further thought to lead to more success in task performance. These findings suggest that motor resonance has close links with social cognition.

Motor resonance could provide a useful tool to investigate the development and integration of new actions and how actions and their verbal referents are learned. It is consequently particularly intriguing to investigate how motor system develop for actions that are currently outside a child's motor repertoire.

1.2.3 Methodological Issues

Before we outline our objectives, it is essential to discuss at this point the difficulties that researchers are facing when conducting EEG research. These difficulties are concerned around data acquisition, analyses and subsequently interpretability as a consequence. A general challenge for both behavioural and psychophysiological developmental research is to adapt adult paradigms and procedures for use with a developmental population. When focusing on neurophysiology, already at the stages of data acquisition, inconsistencies in the literature derive based on a lack of replication in the field and the different methods used across development. This is also seen via an increase in reporting of methods in more recent studies, when contrasted to older

papers. The issues raised in this thesis link to studies from different periods over the past decades. The differences between EEG and ERPs create added complexities in terms of indexing perceptual and cognitive functions. For instance, differences are noted between the infant N400 and the adult N400 even in studies where the same paradigm was applied. Reid et al. (2009) report N400 differences in latency such that the N400 appeared later for the developmental sample when compared to the adult sample. Furthermore, the observed N400 effect was predominantly posterior within the infant sample. The topological difference was explained with the presence of a second overarching component, thought to reflect attentional processes in early development. This component was prevalent in fronto-central region resulting in a potential masking of the N400 component over central sites. Latency differences are often described in developmental ERP research (see Hoehl & Wahl, 2012) and incomplete myelination in the infant brain may be a contributing factor for slower neural responses (Picton & Taylor, 2007). More recently Stets and Reid (2011) discussed how amplitudes change over the course of an experimental session. These changes are not typically discussed within the infant literature but should be considered as they may explain differences in infant and adult responses (Ní Choisdealbha & Reid, 2014). Within EEG work, event-related oscillations are comparably unexplored, however a rapid increase in this field can now be observed – as can be seen in the alpha frequency literature. To this point, developmental studies have not considered a systematic subdivision of this frequency into upper and lower bands as reported in adults. These subbands are thought to reflect different functional properties which are currently being overlooked in developmental research efforts. The use of narrower frequency bands has however the potential to reduce the possibility that frequency specific effects are being missed (e.g., Klimesch, 1999).

Another difficult area in infant EEG is that of data editing and analysis. Researchers are required to find a good balance between guiding their analysis approach on a priori regions or time windows of interest, that are in addition not always consistent within the field, and satisfy the need to take a reasonable account of the data as it presents itself. Reasons for this could stem from a large variety of brain imaging methods used for adult and infant studies which make it difficult to signify which channels on the differing sensor nets correspond to the ERP component of interest. Unfortunately, various methodological issues can obstruct reported findings leading to misconceptions across the field. Expertise in EEG practical techniques, such as the correct application of the EEG net, for instance, is also a considerable factor very rarely discussed in the field. One might argue that the application of the EEG nets per se represents a margin of error. In other words, where the application of the net on the infant scalp is already incorrectly placed, reported effects e.g. from so called frontal sites, could in reality be central just by misplacing the net by a few increments. The choices made during the data editing process are further influenced by the experience of the researcher as well as the standards set by their laboratories. These standards can largely vary from using certain algorithms for trial by trial analysis that can further be followed up with a second round of manual rejection. In addition, the child's behaviour during the testing session is not always assessed. In other words, whether the child has been paying attention to the stimulus material presented is not always considered and consequently trials can enter the final analysis, which are not a direct neurological response to the test stimuli.

For the present thesis, we have adopted current as well as the highest standards in the field, and as a result we believe that we have made appropriate choices based on given literature and our current dataset regarding the choice of channels, time windows

of interest and thresholds for trials entering the final analysis. If we would have taken a more disingenuous approach by including channels based on previous findings, we would have not accurately reflected the data as it presents itself. Further, we adapted new strategies by ignoring fixed band analyses and consequently implemented a narrower frequency band approach. Lastly, each experimental chapter highlights the minimum number of artifact-free trials for participant inclusion and justification for the defined minimum.

We recommend that future infant EEG research should resolve these methodological concerns by introducing standardised protocols. In addition, researchers must report specific details in greater detail related to their procedures and analysis approaches undertaken. In this thesis, we have attempted to report all methodological decisions and approaches as fully as possible

1.3 Thesis Objectives

To date, comparably little is known about the relation between language and action and how semantics is involved within this process in early development. Our primary focus is therefore on the access of semantic meaning via different routes via the help of neurophysiological measures. Three targeted studies were designed in order to better understand the modulation of semantics at distinct points in development.

In Chapter Two, I will be examining semantic processing abilities of everyday non-communicative actions by pre-verbal infants and explore how this relates to language proficiency. My first study utilizes event-related potential methodologies (ERP's) to investigate how social-cognitive processes in pre-verbal infants relate to later language capacities. Nine months old infants' understanding of the semantic structure of actions was assessed via an N400 ERP response to action sequences that

contained expected and unexpected action outcomes. At 9 and 18 months, infants' language abilities were measured using the Swedish Early Communicative Development Inventory (SECDI). We propose that infants who show a selective N400 response to unexpected action outcomes are those who are better language comprehenders and producers. The results will help us explore the notion that language performance is related to the ability to detect and interpret human actions at 9 months of age. They will also further our understanding related to the basic cognitive mechanisms involved in the processing of sequential events that are shared between two conceptually different cognitive domains.

In Chapter Three, semantic representations to newly acquired nouns in 10-11-month old infants were investigated. For this purpose, a combined event-related oscillation (ERO) and ERP study was developed in which infants first experienced a live social interaction for 6 minutes while an experimenter introduced and labeled two novel objects. Immediately after learning and with a 24-hour delay, infants were first primed with one of the two novel words from the learning phase via loudspeakers. Shortly after the word presentation, one of the two objects seen in the learning phase appeared from behind an occluder on a computer screen. The word–object pairing was either congruent or incongruent. By subsequently presenting infants with the objects from the learning phase, that were either in line with the infants' referential expectation (congruent condition) or not (incongruent condition) we postulated that mismatches (incongruent condition) will be detected for the word-object pairs. Specifically, we predicted a more negative N400 component to the incongruent word-object pairs as well as suppression of the upper alpha band in response to congruent word-object couplings. Further, a correlation of upper alpha band suppression with the N400 ERP component would suggest that this specific alpha band is a genuine signature of constructing the

knowledge system at 10- to 11-months. Theta band synchronisation as a representative neural marker for encoding was further considered in this paradigm. Its relation to the formation of new information during infancy was also investigated. Finally, we aimed to contribute to our understanding of semantic formation trajectory by means of infants' memory recall for novel words and objects at two points in time (immediately after and 24-hours later).

The aim of Chapter Four, containing the third experimental study of this thesis, was to investigate the early link between language and motor systems in 24-month-old children. We were specifically interested in examining the modulation of the sensorimotor alpha frequency band before and after the acquisition of novel hand actions, verbs and sounds. We proposed that this rarely performed pre-post training paradigm would give new insight into the relation between motor activation and language acquisition - specifically that of verb acquisition. Our study therefore comprised a pre-training assessment, a 6-day observational home training phase, and a post-training assessment with monolingual English children. We measured verb comprehension as indexed by motor activation of the mu rhythm over frontal and central sites. Further, we investigated whether mu desynchronization in response to action-related verbal material can be distinguished from desynchronization triggered by non-linguistic but equally familiar material. This final manipulation was aimed to explore the notion that verb processing distinctly modulates motor activation in early toddlerhood. Effects would inform us of the functional involvement of both action and language systems.

Chapter 2

Semantic processing of actions at 9 months is linked to language proficiency at 9 and 18 months

Text as it appears in Kaduk, K., Bakker, M., Juvrud, J., Gredebäck, G., Westermann, G., Lunn, J., & Reid, V.M. (2016). Semantic processing of actions at 9 months is linked to language proficiency at 9 and 18 months. *Journal of Experimental Child Psychology*, *Volume 151*, 96-108. Special Issue: Interrelations Between Non-Linguistic and Linguistic Representations of Cognition and Action in Development

Abstract

The current study uses event-related potential methodologies to investigate how social–cognitive processes in preverbal infants relate to language performance. We assessed 9-month-olds' understanding of the semantic structure of actions via an N400 event-related potential (ERP) response to action sequences that contained expected and unexpected outcomes. At 9 and 18 months of age, infants' language abilities were measured using the Swedish Early Communicative Development Inventory (SECDI). Here we show that 9-month-olds' understanding of the semantic structure of actions, evidenced in an N400 ERP response to action sequences with unexpected outcomes, is related to language comprehension scores at 9 months and is related to language production scores at 18 months of age. Infants who showed a selective N400 response to unexpected action outcomes are those who are classed as above mean in their language proficiency. The results provide evidence that language performance is related to the ability to detect and interpret human actions at 9 months of age. This study suggests that some basic cognitive mechanisms are involved in the processing of sequential events that are shared between two conceptually different cognitive domains and that pre-linguistic social understanding skills and language proficiency are linked to one another.

2.1 Introduction

Young infants show sophisticated abilities across an array of cognitive domains during early development (Mandler, 2006). One example of this is the ability to determine outcomes of actions, which is linked to the capacity to process other people's goals and their intentions (Baldwin, Baird, Saylor, & Clark, 2001). It is possible to conceptually divide human action into two primary forms. On the one hand, actions can be communicative in nature and are designed to directly engender social understanding of information. Within this framework, infants' preverbal social–communicative capacities such as the ability to follow others' eye gaze (e.g., Brooks and Meltzoff, 2005, Brooks and Meltzoff, 2008 and Brooks and Meltzoff, 2015) and to process pointing (e.g., Brooks and Meltzoff, 2008, Butterworth and Morissette, 1996 and Tomasello et al., 2007) and gestures (e.g., Kraljević et al., 2014, Rowe and Goldwin-Meadow, 2009a and Rowe and Goldwin-Meadow, 2009b) are well studied and have been related to various aspects of later language abilities.

A second form of actions incorporates all of those actions that are non-communicative in nature but still convey goal directedness. Such goal-directed actions can be observed repeatedly in typical environments by infants during early development, ranging from parents cleaning the home to the preparation and consumption of food. In such scenarios, adults as well as young children readily construct action representations that are organized with respect to this ultimate goal (Baldwin et al., 2001 and Zacks et al., 2001). For example, infants from 6 months of age can accurately predict that a cup (Hunnius & Bekkering, 2010) or a spoon (Kochukhova and Gredebäck, 2010 and Reid et al., 2009) should go toward a person's mouth rather than toward a person's ear. These representations allow for the prediction of the consequences of actions, including the ability to interpret and describe actions and

categorize action sequences (Sommerville & Woodward, 2005). Prior research has shown that the structure of actions parallels that of linguistic utterances and that both actions and language show comparable hierarchical structures (Baldwin et al., 2001 and Zacks et al., 2001). It has been argued that human language emerged from the hierarchical structure of instrumental actions. The same neural circuits that control the hierarchy of these instrumental actions served as a basis on which the newly acquired function of language syntax has emerged (Gallese, 2007). The parallels that can be seen between the semantic organization of non-communicative actions and the semantics within linguistic structures raise the possibility that processing of this action type during early development may be related to language. The context that is present within the execution of sentences and actions conveys information that facilitates the prediction of future events. Situational knowledge thereby provides infants with a mechanism to use semantic rules during action observation (Ní Choisdealbha & Reid, 2014). Among other cognitive and social advantages, this enables infants to reenact the final goal of a modeled action (Gergely, Bekkering, & Kiraly, 2002) and to infer goals of an uncompleted action without seeing the achievement of the goal itself (Daum, Prinz, & Aschersleben, 2008). Given these parallels in the organization of non-communicative actions and the structure of language, it has been conjectured that the ability to process these hierarchically structured actions during early development may pave the way for language acquisition (Reid et al., 2009). From the evolutionary account, it has been suggested that language has its origin in the ability to interpret others' gestures and actions (e.g., Corballis, 2003 and Rizzolatti and Arbib, 1998). Together, these findings indicate that in addition to previous findings, which demonstrated strong relations between infants' preverbal social–communicative capacities and language and have already been well studied, understanding of non-communicative action may also be

linked to the development of language or share similar domain-general processes. To date, there is no empirical evidence to support the notion that links between the two domains exist during early development. Consequently, the current study investigated the relation between the semantic processing of non-communicative instrumental actions during infancy, that is, the ability to detect and interpret others' action end states as either expected or unexpected and relate this ability to language abilities during the first and second postnatal years.

In language research, the N400 component of the event-related potential (ERP) has been identified as a neural signature related to the formation of a semantic representation because the N400 is elicited when a word does not fit an expected context (e.g., Kutas & Hillyard, 1980). In the action domain, N400 effects are observable when action outcomes are violated in infants and adults (e.g., Parise and Csibra, 2012, Reid and Striano, 2008, Reid et al., 2009 and Wu and Coulson, 2005). In adults, the N400 has also been found to be sensitive to the relation between gesture and speech (Holle and Gunter, 2007 and Kelly et al., 2007). The similarity in the electrophysiological responses to semantic violations in the action and language domain suggests that language processing may derive from understanding action during early development or, to some extent, share similar cognitive mechanisms. Should infants first start to understand the parameters of semantic structures within the action domain, this capacity may well bootstrap learning about language. Pre-linguistic semantic processing of action sequences contains hierarchically aligned structures that are similar to that of sentential structures. As such, the capacity to semantically process action may be related to language capacities.

A key question that has to date remained unanswered is whether language capacities are based on structures initially detected and interpreted within the action

domain. For this purpose, we presented 28 9-month-old infants with a sequence of images with expected and unexpected action outcomes known to reliably induce an N400-like ERP component over parietal regions (Reid et al., 2009). In addition to the N400 effect, Reid and colleagues (2009) found a negative central (Nc) component, which was larger in amplitude for the expected condition over frontal and central areas when contrasted with the unexpected condition. The Nc is thought to reflect attentional processes with greater amplitude for stimuli that elicit a higher allocation of attentional resources (Reynolds & Richards, 2005) and is typically found over frontal and central sites, peaking at around 300 to 700 ms after stimulus onset (Webb, Long, & Nelson, 2005). Previously, the Nc has been found to be more enhanced for familiar stimuli, for instance, familiar versus novel faces (e.g., deHaan & Nelson, 1997). Reid and colleagues (2009) argued that their observed Nc effect was driven by infants' higher interest in stimuli depicting food consumption. An alternative and equally plausible explanation given by the authors could be that 9-month-olds are capable of judging where food should be placed when people hold and direct food toward the head area because these are actions that infants are more likely to be exposed to in everyday life. Therefore, a familiarity effect for the expected action conclusion could drive the observed Nc effect. Our study employed the same stimuli and age group as Reid and colleagues' study. Therefore, we expected to replicate the morphology of the Nc and N400 in Reid and colleagues' work within the current study.

To investigate expected and unexpected goal-directed action processing at 9 months of age and infants' language abilities, we assessed language skills at 9 months as well as when the same children reached 18 months of age by introducing two forms of the Swedish Early Communicative Development Inventory (SECDI; Eriksson, Westerlund, & Berglund, 2002). To assess infants' action processing abilities and relate

these to language proficiency, percentile scores from both the SECDI words and gestures (w&g) at 9 months and SECDI words and sentences (w&s) at 18 months were used to split our sample at the 50th percentile (for a similar paradigm, see Torkildsen et al., 2009). Consequently, we examined infants' semantic action processing abilities at 9 months of age (via obtained ERP data) separately for each of the two language proficiency groups and time points (here termed low and high language comprehension at 9 months and low and high language production at 18 months). We hypothesized that if the application of rules relating to the structure of action were also used when processing language structures, only the group of infants with an above mean performance in language comprehension and production at both time points of the language assessment will display evidence for semantic processing within the action domain, as indexed by the N400 ERP component.

2.2 Method

2.2.1 Participants

The final sample comprised 28 9-month-old monolingual Swedish infants (14 female) with a mean age of 277 days (range = 262–287). When split at the 50th percentile on the SECDI at 18 months of age, the low language production group (n = 14, 6 female) had an age range at 9 months of age between 263 and 284 days (M = 272) and the high language production group (n = 14, 8 female) had an age range between 262 and 287 days (M = 276). The average age during the follow-up language production assessment at 18 months for the low language group was 556 days (range = 540–573) and for the high language group was 557 days (range = 547–573). An additional 25 infants were tested but excluded due to inattentiveness, thereby providing fewer than the required minimum number of 5 artifact-free trials per condition for data analysis. A

meta-analysis on 149 published infant EEG studies demonstrated an average attrition rate of 49.2% (Stets, Stahl, & Reid, 2012). Thus, an attrition rate of 47.2% as present in this study is within norms in this research field. Furthermore, 11 infants were excluded due to technical failure ($n = 6$) or parents failing to complete the follow-up language questionnaire ($n = 5$). All infants were born full-term (37–42 weeks) and were of normal birth weight (>2500 g). Each participant's parent gave informed written consent and received a gift voucher for participation. This research was approved by the local ethical committee, which stands in accordance with the 1964 Declaration of Helsinki.

2.2.2 Experimental stimuli

Stimuli were the same as those used by Reid and colleagues (2009). They were composed of three photographs depicting a male or female actor showing eating actions by holding either a spoon, a fork, or the food itself. The sequence of the action followed a hierarchical order always first depicting the action context (A), mid action (B), and action conclusion that either conformed (C) or violated (D) the expected action outcome (Figure 2.1)

Figure 2.1: *ERP stimuli: Three-image sequence depicting context (A), mid action (B), and action outcome (C/D). The N400 and Nc components were assessed for the last image of the sequence depicting the expected (C) or unexpected (D) action outcome.*

Stimuli were shown on a 17-inch computer monitor (37 × 31 cm) at a viewing distance of 90 cm via E-Prime 2.0, E-Studio (Psychology Software Tools, Pittsburgh, PA, USA), and were 22 × 17.5 cm in size. This produced a visual angle of 15.2 degrees (horizontal) and 12.7 degrees (vertical). Images A and B were presented for a duration of 750 ms and Image C/D for 1000 ms, and all three images were presented consecutively. The eight possible three-image sequences were presented at random with the constraint that neither the same condition (congruent or incongruent) nor the same actor (male or female) was presented more than two times consecutively. The stimuli described in the current study were intermixed with other visual stimuli not relevant to this study (a hand displaying a "give me" gesture and a control hand configuration). Both stimulus sets were presented in separate blocks. Each block varied in length ranging from three to eight presentations. The two blocks were presented one after the other throughout the

experimental session to ensure an equal frequency distribution, and the starting order was counterbalanced between participants. To eliminate expectancy effects, the interstimulus interval (ISI) between blocks, as well as between each three-image sequence and the stimuli from the second block, was always presented for a random duration of 300 to 600 ms and in the center of the monitor on a gray background. Each stimulus and block was followed by a black fixation cross on the same gray background, which again lasted for a random duration of 300 to 600 ms. The N400 and Nc components were assessed for the last image of the three-image sequence depicting the (expected or unexpected) action outcome.

2.2.3 Language measure

The SECDI (Eriksson et al., 2002) is the Swedish adaptation of the MacArthur Communicative Development Inventory (Fenson et al., 1993 and Fenson et al., 1994), measuring communicative comprehension and production skills in young children based on parental reports. Two versions are available and are designed for use with different age groups. The SECDI words and gestures is primarily designed to assess language comprehension of 8- to 16-month-olds, and the SECDI words and sentences is designed for 16- to 28-month-olds, which measures only language production. We used the SECDI w&g at the first testing point (9 months) and used the SECDI w&s to investigate infants' language production abilities at 18 months. The SECDI w&g is normed on 228 Swedish children and the SECDI w&s is normed on 336 Swedish children and further provides monthly norms by means of interpolation from existing data (see Berglund & Eriksson, 2000).

2.2.4 Procedure

During the ERP session at 9 months of age, each infant was seated on the caregiver's lap in a dimly lit experimental room. The experimenter was present in the same room but was separated from the parent and infant by a curtain, thereby allowing the experimenter to monitor the infant's behavior via a control monitor. This enabled the experimenter to pause the experiment and offer breaks each time the infant became inattentive or fussy. For the duration of the pause, the monitor remained blank. The experiment ended once the infant's attention could no longer be maintained but did not exceed 10 min in total. The infant was video-recorded throughout the session for offline coding of looking behavior to the experimental stimuli and to control for the amount of motor activity in each trial. After the ERP session, caregivers received the SECDI w&g together with a prepaid envelope and were asked to fill out and return the language questionnaire within 1 week of their first visit by mail.

Participants were followed up 9 months later by contacting the caregivers a week before their infants were due to become 18 months of age. Caregivers were asked to fill out the SECDI w&s online via a link that was provided by the experimenter. Percentile scores from the SECDI w&g and w&s (Berglund & Eriksson, 2000) language measures at 9 and 18 months were used to assign infants to one of two groups, namely low language or high language. In line with previous research (e.g., Torkildsen et al., 2009), we defined the border that separated the low language group from the high language group at the 50th percentile. According to the SECDI manual (Eriksson et al., 2002), the average number of words comprehended by Swedish 9-month-olds is 11. At 18 months, the reported word production reaches 35 on average. Children in the high language comprehension group at 9 months understood a mean of 34.5 words (range = 15–102), and children in the low language comprehension group at 9 months

had a mean total comprehensive vocabulary of 5.6 words (range = 0–11). When followed up, children in the high language production group at 18 months had a mean of 163.4 words (range = 42–484), and the low language production group had a mean total productive vocabulary of 20.4 words (range = 6–33). An independent samples *t*-test confirmed that language scores were significantly higher for infants from the high language comprehension group at 9 months ($M = 74.0$, $SD = 14.0$) than the low language comprehension group at 9 months ($M = 29.2$, $SD = 15.1$), $t(25) = 7.99$, $p < .001$, $d = 3.08$, as well as for the high language production group at 18 months ($M = 82.9$, $SD = 9.9$) in comparison with the low language production group at 18 months ($M = 34.3$, $SD = 14.5$), $t(26) = 10.30$, $p < .001$, $d = 3.90$, thereby justifying the split at the 50th percentile.

However, the two obtained language measures (9-months comprehension and 18-months production vocabulary) did not correlate with one another, $r(26) = .21$, $p = .19$. This is due to the fact that only half of the sample ($n = 14$, 8 high language group and 6 low language group) remained within the same language proficiency group for both time points. That is, 6 children moved up (from low language comprehension at 9 months to high language production at 18 months), and 7 children moved down in their language proficiency (from high comprehension at 9 months to low production at 18 months). For 1 child, we were able to obtain language scores only for the second time point. Consequently, this child was excluded from any analysis concerning 9-months language proficiency comparisons. Thus, children were equally likely to remain the same, move up, or move down in their language proficiency.

It should be noted that word production is preferentially used to mark toddlers' language skills because productive vocabulary is likely to be a more reliable indicator when relying on parental reports to determine language proficiency (Rämä, Sirri, &

Serres, 2013). Language production, therefore, is more commonly used in studies investigating links between language and the N400 (e.g., Friedrich and Friederici, 2006, Friedrich and Friederici, 2010, Mills et al., 2005, Torkildsen et al., 2006, Torkildsen et al., 2008 and Torkildsen et al., 2009). Our a priori main focus, therefore, was to assess the relation between word production and the action N400.

2.2.5 EEG recording and analysis

Electroencephalography (EEG) was acquired via 128-channel Geodesic Sensor Nets amplified with an EGI Net Amps 300 amplifier with a sampling rate of 250 Hz (HCGSN 130, EGI, Eugene, OR, USA). EEG was further referenced online to the vertex electrode (Cz), and an analogue hardware bandpass filter of 0.1 to 100 Hz was applied with data stored for off-line analysis. Subsequent data editing was performed using Net Station 4.5.4 Waveform Tools and included digital filtering (0.3–30 Hz bandpass, Kaiser type Finite Impulse Response filter, passband gain 99.0%–0.1 dB, stopband gain 1.0%–40.0 dB, filter roll-off 0.29 Hz) and data segmentation (200 ms before and 1000 ms after stimulus onset). The 200-ms baseline used for segmentation represented the last 200 ms of the mid action (B) image (for the same procedure, see Reid et al., 2009). Off-line inspection of the video-recordings ensured that only trials in which infants paid attention to all three images of the given sequence were further processed. The remaining trials were manually and visually inspected for artifacts caused by eye or body movements using normative procedures for infant ERP studies[3] (see Hoehl & Wahl, 2012, for further information on infant ERP methods). An artifact detection tool was applied, and EEG data in affected channels were marked as bad whenever amplitude exceeded the set threshold of ±200 μV for the entire segment, with

[3] The EEG pre-processing was not done blind to condition.

a moving average of 80 ms. Trials including more than 10 channels with excessive noise levels were rejected. Noisy channels from the remaining trials, which comprised fewer than 10 bad channels, were interpolated by using the average of surrounding electrodes. Remaining artifact-free trials were re-referenced to the average electrode and baseline corrected before creating individual averages for all participants. The grand average was produced from those 28 individuals who contributed 5 trials to their average per condition[4]. Because the second image also serves as the baseline for the third image, we did not allow for any shifts in attention throughout the three-image sequence given that this would have been reflected in a noisy baseline. On average, infants saw 34.2 three-image sequences (range = 20–48) for the expected condition and 34.3 (range = 21–49) for the unexpected condition. The probability of attending to all three images of the sequence decreased with the length of the testing session. Based on findings from Stets and Reid (2011) as well as Kaduk, Elsner, and Reid (2013), reliable and interpretable data have previously been shown to be present within infant EEG with as few as 3 to 7 trials per condition using standard visual ERP paradigms. The Nc component throughout the current analysis, therefore, is useful as a guide to indicate the morphological quality of the current dataset. A very commonly observed problem, besides infants' general lack of attention in ERP studies, is neural habituation. Neural habituation potentially alters the outcome of obtained ERP morphology. Stets and Reid (2011; see also, e.g., Snyder, Webb, & Nelson, 2002) reported that infants' habituation to stimuli causes a shift of attention over the course of an experimental session. The inclusion of trials obtained nearer the end of an acquisition session can attenuate earlier

[4] The choice of the first five good trials was not decided prior to the study data collection, although it was implemented prior to analysis of the data. The study was collected at the same time as another EEG study (Bakker et al., 2015) and as such, the number of trials acquired for each condition was seen to be relatively reduced before data analysis. As recent work (e.g., Stets & Reid, 2011) showed that meaningful morphologies could be examined with relatively few trials, this analysis approach was taken from the outset with this dataset.

effects with the potential to even produce opposite effects. Infants from the current study largely differed in the amount of trials they were able to contribute to their individual averages between participants, ranging between 5 and 20 trials for the expected condition and between 5 and 21 trials for the unexpected condition. Including "as many trials as possible" would potentially mask any Nc effect if taking into account the changes in amplitude through neural habituation. Therefore, rather than including as many artifact-free trials as possible for each participant's individual average and condition, we decided to analyze only the first 5 artifact-free trials[5]. Furthermore, the inclusion of a larger sample of infants compensates for the low inclusion criterion. Because there is currently no evidence that a larger number of trials improves the morphology of the N400 component during early development, we believed that by replicating the Nc effect found in Reid and colleagues (2009) we would further provide evidence for the validity of this approach and would justify the employment of this technique with the N400[6].

2.3 Results

2.3.1 Nc ERP component

Visual inspection of the Nc effect led to the identification of 10 channels representing the left central scalp (29, 30-C1, 35, 36-C3, 37, 40, 41-C5, 42, 46, and 47) and 10 channels representing the right central scalp (87, 93, 98, 102, 103-C6, 104-C4, 105-C2, 109, 110, and 111), which were in line with topographical locations for the Nc present in prior infant research (e.g., Parise, Reid, Stets, & Striano, 2008). A time window was

[5] Analysis was always based on five trials per condition for each participant. There was therefore no difference between conditions.

[6] The Nc and the N400 were defined by utilising the same topographies as in prior research. Given the difference in arrays between prior research and the large number of electrodes in the present study, the region of the analysis was nonetheless similar. With this dataset, a data driven analysis was not required as there was an alignment between prior studies and the current dataset for the topography and morphologies involved.

chosen from 350 to 600 ms post-stimulus presentation in order to capture the primary peak of the component.

For the overall sample ($N = 28$), we first performed a 2 × 2 repeated measures analysis of variance (ANOVA) with the within-participant factors condition (expected or unexpected) and hemisphere (left or right). Statistical analysis revealed a significant main effect for condition, $F(1, 27) = 4.37$, $p = .046$, partial $\eta^2 = .139$ (see Figure 2.2).

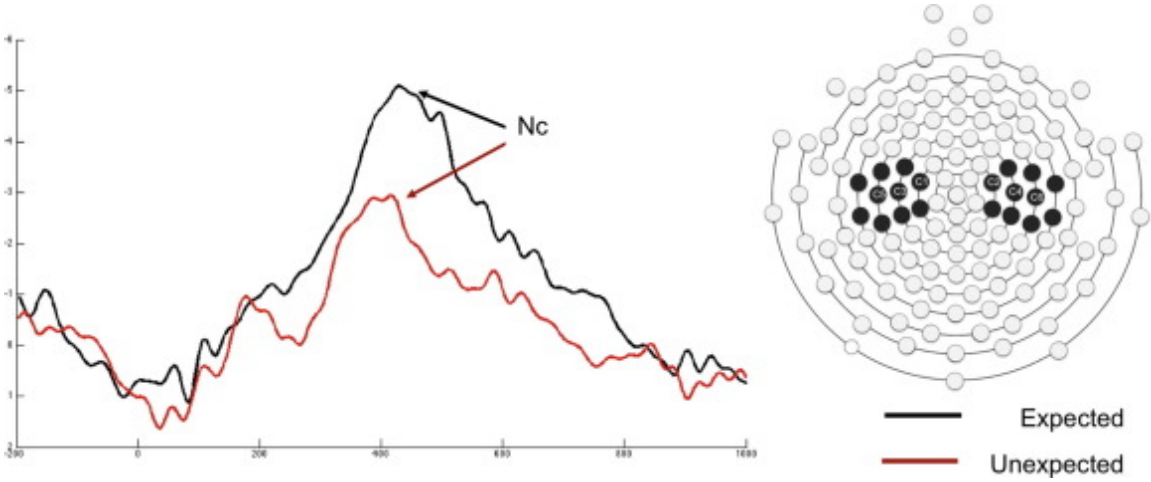

Figure 2.2: *Left: Averaged ERPs for all infants (N = 28) from 20 channels, 10 of the left central region (29, 30-C1, 35, 36-C3, 37, 40, 41-C5, 42, 46, and 47) and 10 of the right central region (87, 93, 98, 102, 103-C6, 104-C4, 105-C2, 109, 110, and 111), displaying the Nc. Right: Analyzed area of interest with reference to the 10–20 system.*

2.3.2a Nc: Language division from the SECDI w&g for language comprehension

When dividing the sample into groups according to their language comprehension abilities at 9 months of age, the 2 × 2 repeated measures ANOVA with the within-participant factors condition (expected or unexpected) and hemisphere (left or right) and the between-participant factor language comprehension revealed a significant main effect for condition, $F(1, 26) = 4.45$, $p = .045$, partial $\eta^2 = .151$ ($M_{expected} = -3.83$, $SD = 4.05$, $M_{unexpected} = -1.93$, $SD = 3.51$). No other main effects or interactions were found (all p values > .05).

2.3.2b Nc: Language division from the SECDI w&s for language production

For the groups split according to the SECDI w&s at 18 months of age, via which language production was measured, the 2 (Condition) × 2 (Hemisphere) repeated measures ANOVA with the between-participant factor language production revealed a significant main effect for condition, $F(1, 27) = 4.35$, $p = .047$, partial $\eta^2 = .143$ ($M_{expected} = -3.62$, $SD = 4.13$, $M_{unexpected} = -1.79$, $SD = 3.53$). No other main effects or interactions were found (all p values > .05).

A follow-up paired sample t-test comparing both groups (high and low producers) for the central electrodes confirmed these results given that no significant differences between conditions were found in the high language production group, $t(13) = -1.985$, $p = .069$, $d = 0.69$ ($M_{expected} = -3.72$, $SD = 3.61$, $M_{unexpected} = -1.07$, $SD = 4.08$) or with the low language producers, $t(13) = -0.881$, $p = .394$, $d = 0.26$ ($M_{expected} = -3.51$, $SD = 4.70$, $M_{unexpected} = -2.51$, $SD = 2.85$).

2.3.3 N400 ERP component

For the analysis of the N400, a time window around 600 to 800 ms after stimulus onset in parietal sites was chosen based on prior literature (Reid et al., 2009) and the observed ERP morphology because the current research is the first to report on N400 to action with this population using a high-density EEG array. Based on visual inspection of the individual and grand averages, we chose 16 electrodes to represent the parietal region (52-P3, 53, 58, 59, 60, 61, 62-Pz, 67, 72, 77, 78, 85, 86, 91, 92-P4, and 96). In line with prior N400 infant ERP data (Reid et al., 2009), a peak effect was observed in the unexpected condition but not in the expected condition (Figure 2.3)[7].

[7] The large negativity at the end of the trial was due to the offset of the visual stimulus. This can be seen from 950ms.

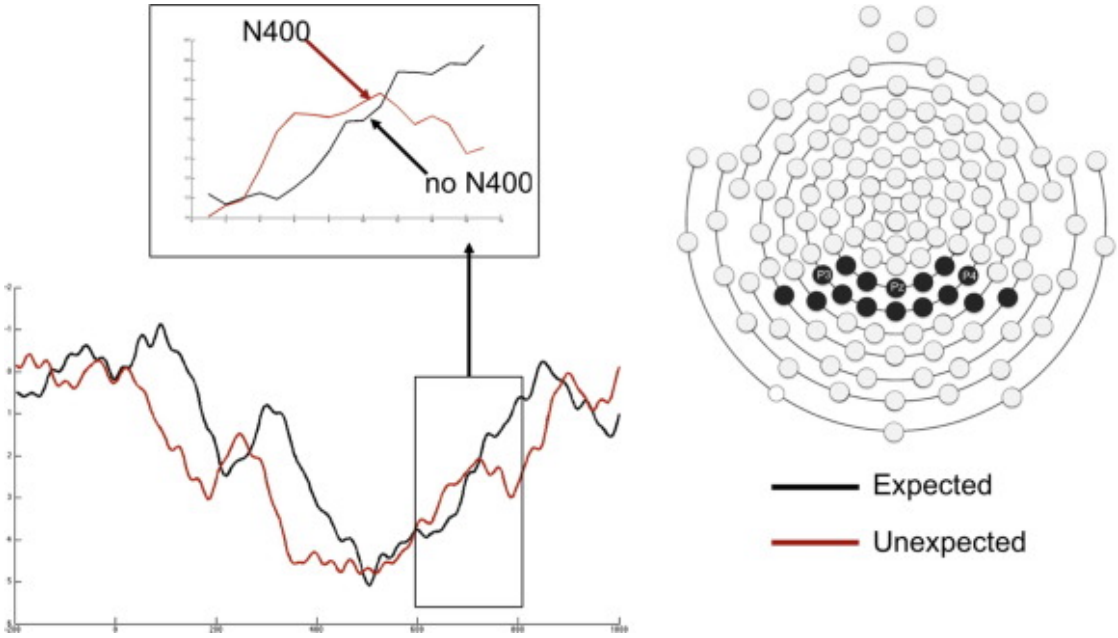

Figure 2.3: *Bottom left: Averaged ERPs for all infants (N = 28) from 16 channels within the parietal region for the N400 (52-P3, 53, 58, 59, 60, 61, 62-Pz, 67, 72, 77, 78, 85, 86, 91, 92-P4, and 96). Top right: Zoomed-in time window of interest indicating the N400 for the unexpected condition and the no N400 for the expected condition. Right: Analyzed area of interest with reference to the 10–20 system.*

The data were consequently analyzed using a time window analysis technique (Hoorman, Falkenstein, Schwarzenau, & Hohnsbein, 1998), which we performed overall as well as separately for each group (low language and high language) to assess the morphology of the N400 component within each of the groups. Time window analysis is used when one condition displays a defined peak in the window of interest, whereas the other does not. Therefore, time is included as an additional within-participant factor in the repeated measures ANOVA. A significant interaction between condition and time indicates validity for the hypothesis that the N400 component is evident only for the unexpected condition as a function of variance over time. The repeated measures ANOVA, therefore, featured the two factors condition (expected or

unexpected) and time (17 samples at 1 sample per 12 ms), as suggested by Hoorman and colleagues (1998) and further applied by Reid and colleagues (2009) and Domínguez-Martínez, Parise, Strandvall, and Reid (2015) in N400 research. We selected 17 samples at 12 ms because this best reflected the difference between conditions. First, to confirm previous findings from Reid and colleagues (2009), we examined significant interactions over the entire sample. Results for the overall sample confirmed the observed morphology of the N400 effect for the unexpected condition but not for the expected condition, with a significant condition and time interaction (Greenhouse–Geisser correction), $F(3.6, 98.0) = 3.273$, $p = .018$, partial $\eta^2 = .108$.

3.3.3a Analysis considerations

However, for the N400 time window analysis (Hoorman et al., 1998), a three-way interaction would be very difficult to establish when there is an unequal amount of variance across the two groups, which is what is predicted and shown from this type of analysis. The variance in the low language group would need to be considerably more homogeneous for a three-way interaction to emerge. In cases like this, when looking directly at variability in developmental data, it is strongly recommended that groups be addressed separately (Thomas et al., 2009). Therefore, we hypothesized that we would find a significant interaction between condition and time exclusively for the high language production group when further dividing the two groups according to language proficiency.

3.3.3b N400: Language division from the SECDI w&g for language comprehension

The results from the 2 (Condition) x 17 (Time) ANOVA with the between-participant factor 9-months comprehension showed a significant interaction between condition and time (Greenhouse–Geisser correction), $F(3.5, 86.8) = 3.043$, $p = .027$, partial $\eta^2 = .109$.

No three-way interaction with the additional between-participant factor language comprehension was evident. We further divided the sample according to language proficiency at 9 months of age. We again performed a 2 × 17 ANOVA and found a significant interaction between condition and time for the high language comprehension group at 9 months, $F(2.5, 34.4) = 3.386$, $p = .037$, partial $\eta^2 = .195$. Infants from the low language comprehension group at the 9-months language split did not produce a significant effect for the condition and time interaction, $F(3.6, 39.3) = 1.182$, $p = .332$, partial $\eta^2 = .097$.

3.3.3c N400: Language division from the SECDI w&s for language production

Results from the 2 (Condition) × 17 (Time) ANOVA with the between-participant factor 18-months production revealed a significant condition and time interaction, $F(3.6, 92.63) = 3.225$, $p = .020$, partial $\eta^2 = .110$. The 2 × 17 ANOVA performed for the two language proficiency groups separately revealed a significant interaction only for the high language proficiency group when split according to the language measure assessed at 18 months, $F(3.5, 46.1) = 4.075$, $p = .009$, partial $\eta^2 = .239$. Infants from the low language production group did not produce a significant effect for the condition and time interaction, $F(2.9, 37.4) = 0.901$, $p = .446$, partial $\eta^2 = .065$. To illustrate the interaction over time, we computed a difference wave topographical map (expected – unexpected) for all four language proficiency groups over the course of 200 ms in the observed time window of interest between 600 and 800 ms (Figure 2.4).

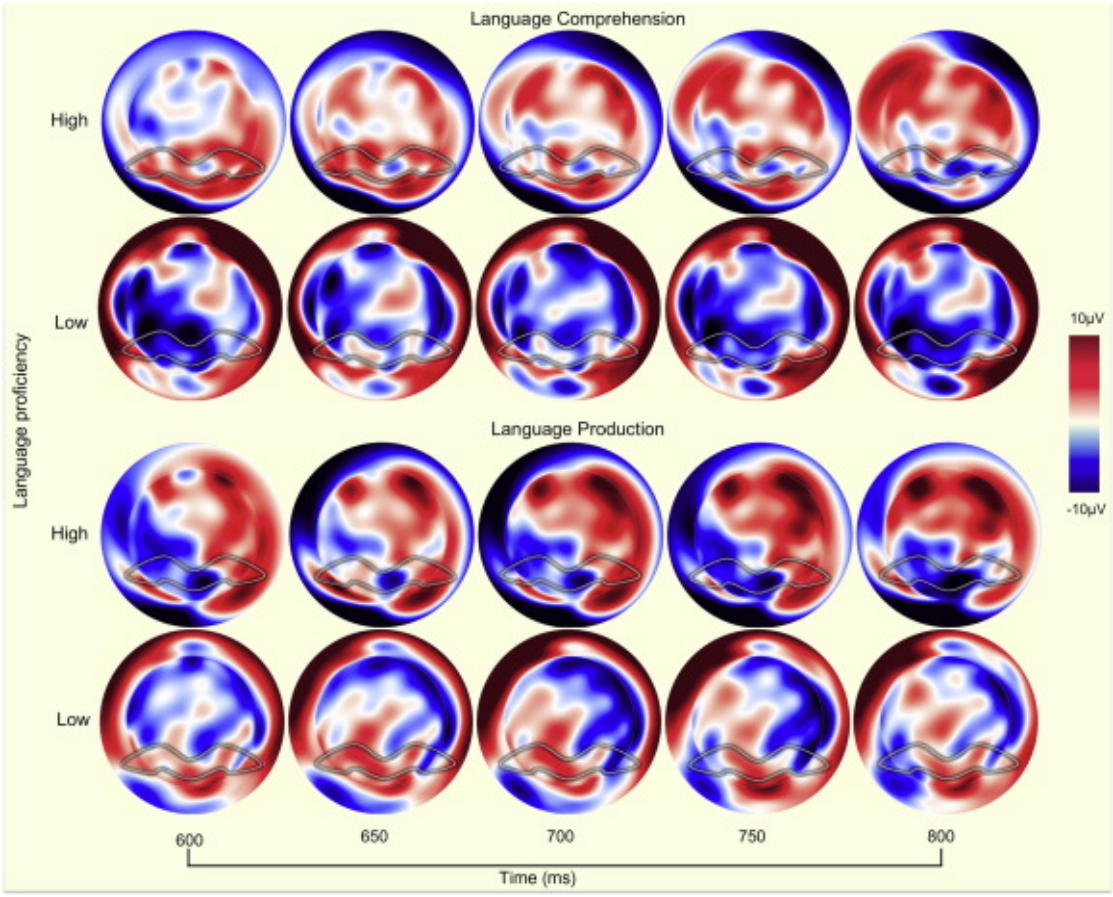

Figure 2.4: *Difference wave (expected – unexpected) topographical maps for high and low language comprehension groups at 9 months of age (top two rows) and high and low language production group at 18 months of age (bottom two rows) displaying every 50 ms of the 200-ms time window of interest. The effect is clearly seen in high language posterior bilateral regions in the highlighted area, with a high intensity increase in microvoltage for the high language group from near the outset of the epoch, which continues for the duration of the epoch. In contrast, the low language group remains static and does not change in topography in terms of differential microvoltage for the entire epoch in the area of interest.*

2.4 Discussion

The current study used event-related potential techniques to investigate how the understanding of non-communicative actions during early development may relate to language capacities later in life. Our results successfully replicated both the N400 and Nc components in related regions of the scalp and time windows as those reported in

Reid and colleagues (2009), with the ERPs following the same morphology despite the use of a high-density electrode array, whereas the original study had less than 30 channels for the overall sample[8].

When looking at the obtained ERP data by dividing the group as a function of language comprehension at 9 months of age, the N400 effect was evident only in the group of children who scored higher on the SECDI w&g. The same pattern of effect remained stable when dividing our ERP results based on language production capacities at 18 months of age. Overall, we can draw the conclusion that the ability to judge whether an observed non-communicative human action is to be expected or not on the semantic level is closely related to language performance. During the processing of expected and unexpected actions, only the high language production group appeared to differentiate the two actions as indexed by the presence of the N400 ERP component in the unexpected action condition.

The overall main effect for condition for the Nc component over the entire sample indicated that more attentional resources were used to process the expected, and arguably more familiar, action outcomes. When the group was divided by language comprehension abilities at 9 months of age, as well as high and low production at 18 months of age, neither language group was able to discriminate between the familiar and unfamiliar action outcomes. The general conclusion that can be drawn from this is that both groups seem to be equally capable of detecting the familiar action outcomes from the unfamiliar ones by allocating more attention to the more familiar action outcomes.

[8] The initial selection of electrodes was driven by prior research. It is a standard approach when examining infant ERP components to determine the topographical distribution via visual examination of the data (Hoehl & Wahl, 2012), although it should be noted that no statistical evidence currently exists that this approach is the most appropriate. Research is being conducted on these forms of methodological issues (Domínguez-Martínez, 2017).

We divided our sample at the 9-months comprehension level as well as at the 18-months language production level. Children in the current study were equally likely to remain within the same language proficiency group as they were to move up or down in their language proficiency. These results confirm the generally perceived weakness across literature toward the reliability of parental reports on the comprehension level (e.g., Friedrich and Friederici, 2006 and Friedrich and Friederici, 2010; Mills et al., 2005; Rämä et al., 2013, Torkildsen et al., 2006, Torkildsen et al., 2008 and Torkildsen et al., 2009). Therefore, it is important to stress that any analyses reported in this study and interpretations that are based on splits at the 9-months comprehension level need to be taken cautiously. Although we were not able to find three-way interactions with language proficiency as the between-participant factor in our performed repeated measures ANOVAs, the targeted statistical analyses reported in the current study mirror prior research investigating the link between language proficiency and the language N400 (e.g., Friedrich & Friederici, 2010). Due to our a priori hypothesis, separate analyses as a function of language capacity were undertaken.

Prior work has shown that aspects of social communication are linked to later language development. The observed effects of the current study occur earlier in ontogeny when compared with other known aspects of social understanding, including periods where the ability to interpret eye gaze (Brooks & Meltzoff, 2005) or gestures (Butterworth and Morissette, 1996, Rowe and Goldwin-Meadow, 2009a, Rowe and Goldwin-Meadow, 2009b and Tomasello et al., 2007) predicts later language proficiency. The current study links with prior language development research but provides a different pathway in the attempt to understand the development of language. Specifically, the current study investigated non-communicative human actions using neurophysiological measures of action semantics reflected in the N400 ERP component.

Here we suggest that language proficiency is aligned with semantic processing abilities during development. The current data suggest that these competencies in action may co-occur or bootstrap learning about language throughout development. When contrasted with previous work investigating early social skills and later language abilities (e.g., Brooks & Meltzoff, 2005 and Brooks and Meltzoff, 2008), the current study introduces an additional measure in the form of a neural marker that indexes semantic understanding of action sequences. Prior literature could not comment on the underlying structure of cognitive processes because all prior literature is based on inference from behavioral results. The ERP approach provides a basis for a stronger conclusion in terms of the specific cognitive activities related to processing action sequences. Rather than indicating that relations exist between the action and language domains, the current study allows for the assessment of a specific cognitive mechanism in the form of the N400. Because the N400 is involved only during semantic tasks, this cognitive process must be involved in the processing of actions. Therefore, it is possible to conclude that the semantic processes that are engaged during action observation are linked to language proficiency. With regard to the stability and predictability of language development, we need to stress that measures of language during infancy and early childhood are thought to have issues with reliability. Specifically, measures of language capacities are thought to become increasingly more consistent later in development (Bornstein & Putnick, 2012). As such, our measures at 9 months of age should be taken with caution.

Other aspects of action prediction not assessed in the current study may well be related to later language skills. For example, the relation between N400 action semantics and N400 language semantics is not yet explored. Further studies are required

to map the parameters of these processes across development by using a more refined measure for language comprehension earlier in development.

This study assessed an electrophysiological component where the participants required no overt behavioral response. The learning mechanisms underlying our findings may be explained via the directed attention model (Reid & Striano, 2007) that has implications for understanding social learning during infancy. In this model, different aspects of information lead to more refined processing of environmental variables. This allows infants to focus limited attentional resources to optimize the outcome of any social interaction. For example, the early detection of semantic structures in actions and the enhanced attentional allocation to familiar action outcomes will allow infants to disambiguate action that may help to determine the referent of words or to understand sentences. The Nc effects reported in the current study align well with such an interpretation of the development of social processing skills during infancy, with attention allocated to specific action outcomes.

2.5 Conclusion

These results provide evidence for a developmental account that focuses on how language function emerges from pre-linguistic social understanding skills over the first postnatal year. Following these results, experiments can now be performed whereby the nature of this relation can be further examined, with the N400 being used as a primary tool to determine the associations between action and language processing during early development. This work highlights the parallels in the organization of actions and language, demonstrating that some basic cognitive mechanisms involved in the processing of sequential events are shared between conceptually different cognitive domains that already coexist early in development. Furthermore, this study raises the

intriguing concept that understanding actions during early development may to some extent evolve at the same time as language acquisition or act as a mediator for linguistic skills. Finally, on the basis of these data, it is clear that observed overall effects may well provide important information related to development when examined via an individual differences approach.

2.6 Authors Contributions

K.K. and V.M.R conceptualized this study and developed the methodology. K.K., M.B., and J.J. were involved in data acquisition and K.K. undertook data editing of the acquired data. K.K, V.M.R., and J.L. performed the formal analysis. The original draft was written by K.K. who also provided visualizations. K.K, V.M.R., G.W., M.B., G.G., J.L., and J.J. were involved in review and editing and resources were kindly provided by G.G.

Prelude to Chapter 3

The Language N400 ERP and a potential additional measure for accessing semantic knowledge

Utilizing ERP measures in the first experiment provided a solid basis for a stronger conclusion in relation to specific cognitive activities related to processing everyday action sequences. We demonstrated that certain rudimentary cognitive mechanisms that are involved when 9-month-olds process action sequences are likely linked with language proficiency. In other words, unexpected action outcomes were processed semantically for the groups of children, who were classed as high in their language proficiency. This was evident in the presence of the action N400 ERP to actions that did not fit the preceding context. It should be noted that children from the low language proficiency group could not judge the same action sequences semantically as reflected in the absence of the N400 ERP. However, we anticipate this capacity to be acquired later in development for this group. One way to validate that infants who do not show an N400 ERP response are not able to interpret action sequences at this stage in development is to design a study whereby the N400 activity can be linked to behavioural measures. Evidence for this approach derives from behavioural research such as that of Hunnius and Bekkering (2010) who showed that infants are able to predict goals of familiar actions. Behavioural results should be reflected in the corresponding neurological activity.

The Nc ERP response did not differ between the two language proficiency groups leading us to exclude attentional mechanisms as a driving force for the detection

of semantics. However, it may also be possible that our findings are driven by those infants that are developmentally more facilitated. In other words, it is possible that our paradigm indexes general cognitive development on the basis of the measures we have taken.

As for the use of language measures, one has to also take into consideration potential issues of language measures more over when it comes to parental judgement of their child's comprehension abilities (e.g., Friedrich and Friederici, 2010; Mills et al., 2005; Rämä et al., 2013, Torkildsen et al., 2006, Torkildsen et al., 2008; Torkildsen et al., 2009). Therefore, inferences based on parental reports for comprehension language measures should be always taken with caution. Within our dataset, we could observe large intra- and inter-individual language variations. These variations may stem from the nature of early language development per se and not solely due to the sensitivity or validity of the applied language measure. This is because in early development, even just a few words that are understood or comprehended can lead to a child's language skills being classified in different ways. Bornstein and Putnick (2012) suggest that measures of language capacities become increasingly more consistent later in development. In addition, we encourage future research to use a more longitudinal approach when studying language development to overcome variability in language development. Hereby we feel that by applying N400 ERP measures in addition to parental reports, we may have the capacity to monitor the time-point in development when children's abilities to detect action sequences are aligned with their language proficiency.

As the N400 is specifically involved during semantic tasks, this cognitive process must be involved in the processing of actions. Assuming a likely link between the action and the language domain as indicated by our findings, we subsequently

wondered if the language induced N400 ERP would show similar patterns of activation as that of the action domain. This would serve as a stronger case to argue for the N400 observed in both cognitive domains to likely reflect the same construct. Hereby it is important to note that language comprehension relies on the involvement of other language processes such as syntactic knowledge. In this context, we need to distinguish the N400 properties in response to violation of semantic and syntactic violation. Work by Gunter and colleagues (1997) already very early on postulates that it is semantic violation that elicits an N400 response, whereas syntactic violation is not.

Parise and Csibra (2012) have previously demonstrated that 9-month-old infants are able to detect mismatches between familiar objects when these objects were initially primed by the auditory language referent. The authors used a less frequent method within developmental language N400 ERP literature for the semantic violation paradigm. This method entails first priming the infant with the word. This represents an intriguing way to investigate semantic knowledge in early development, in a manner unexplored in the action domain. We can conjecture that the language induced N400 should first allow for an initial build up of a mental representation, triggered by auditory stimulation of a word. Purely based on infant's previous semantic knowledge, subsequently presented objects on the screen should therefore either induce the N400 language ERP (word-object mismatch), or not. Motivated by this approach, we hoped that this method would reflect more genuine processes of semantic knowledge recruitment given the absence of visual cues.

However, ERP's alone provide only limited understandings with respect to brain processes. A growing interest in developmental cognitive neuroscience therefore increasingly focuses of event-related-oscillations (ERO's) (e.g. Hoehl, Michel, Reid, Parise, & Striano, 2014). In adults, Klimesch et al. (1996) related the presence of theta

synchronization during word encoding to successful recall and later added that a distinct band within the alpha range (upper alpha desynchronization) reflects activation of the semantic knowledge system (Klimesch, 2012). Frequency analysis techniques may well serve as a useful addition in order to help us understand underlying neural underpinnings of semantic processing. Consequently, the goal of the second experiment was motivated by the paradigm applied by Parise and Csibra (2012) in order to investigate language induced N400 ERP, as well as additional measures for assessing semantic knowledge.

Chapter 3

The Trajectory of Semantic Representation: from Encoding to Consolidation - an interplay of N400 & Alpha Desynchronization

Abstract

The present study combines event-related brain potentials (ERP's) and event-related oscillations (ERO's) to investigate semantic representation of newly acquired nouns at two time points. Twenty 10-11-month-old infants first experienced a 6-minute live social interaction in which an experimenter introduced and labeled two novel objects. We investigated infants' retention for the newly acquired nouns and objects immediately after training and with a 24-hour delay. First infants were primed with the words from the learning phase via loudspeakers. Shortly after, one of the two objects appeared from behind an occluder on a computer screen. The word-object pairs were either congruent or incongruent. Our results suggest that new information transforms to a more long-term representation over time. This was evident via the observed N400 ERP component for incongruent word-object pairs following the 24-hour delay. We further focused on two neural markers thought to index information encoding (theta synchronization) and recruitment of knowledge (upper alpha desynchronization). We found alpha band suppression for congruent word-object pairs over posterior channels. Both alpha and the N400 were only evident in the delayed retention test. Further, a correlation between the N400 and alpha suppression was found, suggesting that this specific alpha frequency recruits the knowledge system at 10- to 11-months. This implies that both correlates may index related cognitive mechanisms, namely the processing of semantics, when applied in semantic expectancy paradigms. In addition, we examined differences in attentional mechanisms between the immediate and delayed retention phases. As with the N400 component, the Nc negative central component (associated with allocation of attention) was greater for incongruous word-object pairs only 24-hours later. Given these consistent findings, this study highlights the importance of offline consolidation and proposes that semantic representation of newly acquired words in memory is most likely a function of time.

3.1 Introduction

One pivotal component of early language acquisition is the ability to learn and remember what a word refers to in the wider world. In the case of noun acquisition, infants first need to form an association between an auditory representation of a word and a visual representation of an object (Borgström, von Koss Torkildsen, & Lindgren, 2016). However, whether infants can form stable word-object associations that reflect referential semantic understanding of speech, rather than merely associations between objects and referents, is currently not fully understood. The literature seems to suggest that preverbal infants may utilize multiple cues (perceptual, social, linguistic) when learning new words and that the social context may inform the learner about semantics (Ní Choisdealbha & Reid, 2014).

Neurophysiological parameters such as event-related potentials (ERP's) have proven to be a useful tool for investigating knowledge acquisition in preverbal infants. In social situations, ERP's indicate that live encounters increase information flow that may facilitate learning. For instance, Parise, Reid, Stets and Striano (2008) designed a live paradigm in which 5-month-olds experienced two blocks, each including a pre-test followed by a post-test where EEG was recorded. In both pre-tests, an experimenter introduced small toys on a computer screen. The experimenter was looking at the infant and then to the objects on the screen in 5-second intervals with a friendly face and tone of voice. The only difference between the two pre-tests was in the presence or absence of mutual gaze, which was described by the authors as a "fundamental component of joint attention". In the post-test, infants were presented with all previously introduced objects at random while EEG was recorded. As indicated by a more negative Nc component, infants allocated more attentional resources to objects when eye contact took place in the pre-test phase. Similar neural patterns to those in Parise et al. (2008)

were observed in Striano, Reid and Hoehl (2006) where an adult again engaged in a live paradigm in two conditions with 9-month old infants. In the joint attention context, the adult gazed at the infants' face first and then to a computer screen, displaying novel objects. In the non-joint attention condition, the adult only gazed at the novel objects on the screen while talking to the infant simultaneously. The results suggested that seeing someone directing attention to an environmental object induces a shift of attention in the infant, resulting in the alignment of both subjects' attention on to the same object, which was reflected in an enhanced Nc component during the processing of objects in the joint attention condition when compared to the non-joint attention condition. These findings indicate that infants profit from joint attention when processing information in their surrounding environment. This leads to suggest that a live joint attention interaction, which is accompanied by eye contact but does not necessarily need to co-occur with the test trials as shown by Parise et al. (2008), affects the processing of objects. Kopp and Lindenberger (2011) investigated joint attention and its relation to short and long-term memory in 9-month-olds using a similar paradigm to Striano et al. (2006). The authors found differences in the Nc component between new and old items to be greater in delayed retention when compared to the immediate retention phase with the Nc eliciting more negativity for new items following a one-week delay. The reported Nc effect may be the result of a greater differential perception between new and old items in the second session. Unlike Parise et al. (2008) the reported findings by Kopp and Lindenberger (2011) were evident regardless of the degree of joint attention during initial familiarization.

The N400 ERP component has been reliably identified as a neural marker for semantic processing in language research (Kutas & Hillyard, 1980) and during action processing (e.g., Domínguez-Martínez et al., 2015; Kaduk et al., 2016; Reid et al, 2009)

from adulthood to infancy. In both cognitive domains, the N400 was found to be a sensitive indicator of the semantic relationship between the target and context. In other words, the N400 component is prone to elicit more negativity when the target does not fit the semantic representation of the context in which the target occurred. Results from infant studies using this neural correlate as a tool to investigate semantic processing are currently not consistent. For instance, Parise and Csibra (2012) demonstrated that 9-month-old infants are able to detect the mismatch between familiar objects appearing from behind an occluder and a preceding label, as indexed by a larger N400 component to the incongruent word-object pairs - only if their own mother, but not an experimenter, introduced those objects in a live paradigm. The authors concluded that word-to-object priming occurs referentially in 9-month olds if a) infants appreciate the referential nature of words conveyed through their mother, leading to an activation of the object features associated with familiar objects and b) if there is no requirement to learn new words.

Research focusing on word acquisition reports that infants as young as 6-month of age are able to match novel words with arbitrary visual referents in a few trials by taking advantage of prosodic cues (Friedrich & Friederici, 2011; Shukla et al., 2011). However, the electrophysiological signs of semantic priming disappeared after 24-hours, suggesting strong limitations in memory processes in the first half of the first postnatal year (Friedrich & Friederici, 2011). By the time infants reach the age of 12 months, N400 effects are missing completely, suggesting insufficient lexical-semantic memory structures (Friedrich & Friederici, 2008). Fourteen-month-old infants are, however, able to learn arbitrary mappings between words and objects within a very short training session, even with the absence of social cues. In addition, the newly acquired knowledge could be recalled for at least one day (Friedrich & Friederici,

2008).

Separately to N400 research, an increasing body of research is proposing a so-called dual memory system, in which integration of new knowledge with long-term memory happens over time (e.g., Henderson, Weighall, Brown, & Gaskell, 2012; Henderson, Weighall, & Gaskell, 2013). Henderson and colleagues (2013) for instance, demonstrated that 5- to 9-year-olds were better in recalling newly acquired vocabulary with a 24-hours delay as opposed to immediately after training. An even more stable lexical representation was built for those children who, in addition, were exposed to the meaning (semantic information condition) of the new words during training (Henderson et al., 2013). Children seem to exhibit similar offline consolidation processes as adults to fully integrate new words into the existing lexicon (Brown, Weighall, Henderson, & Gaskell, 2012). Research focusing on early word learning highlights the importance of sleep for memory consolidation. In the case of children's memory formation, sleep is a particular facilitator for recently-encountered novel words and word meaning (Axelsson, Williams, & Horst, 2016). Effects of offline consolidation (that is, the reorganization of recent memories and the effects of formation of semantic knowledge as a function of sleep), are further discussed in a more recent ERP study by Friedrich, Wilhelm, Born, and Friederici (2014). In this study, 9- and 16-month-old infants who napped during the retention period were able to remember specific word meanings and successfully generalized words to novel categories as indexed by the N400 component over central and parietal midline regions. During the training, the authors reported an absence of the N400, indicating that lexical semantic knowledge was not acquired online during the training session in either the nap or the no-nap group across the two age groups (Friedrich et al., 2014).

In sum, the current literature on semantics and memory processes for familiar and newly acquired words and their occurrence, as a function of time, is contradictory. While some studies report the disappearance of semantic priming after 24-hours (Friedrich & Friederici, 2011) more recent ones suggest that offline consolidation is beneficial for the formation of semantic knowledge (Brown et al., 2012; Friedrich et al., 2014). In Parise and Csibra's (2012) study 9-month-olds word-to-object priming for familiar words occurs referentially and only if the mother introduced the object probes, whereas, on the other hand 6-month olds were able to match novel words with arbitrary visual referents in only a few trials (e.g., Friedrich & Friederici, 2011; Shukla et al., 2011). This highlights the importance for further research and for the assessment of a wider array of neural correlates that index aspects of information processing and cognitive function.

"ERP's provide good insights into brain processes however a growing body of research is also including event-related-oscillations (ERO's) to investigate object processing (e.g. Hoehl, Michel, Reid, Parise, & Striano, 2014). To fully investigate semantics in early development in the current study, we therefore propose to also expand our investigations by considering ERO's. Two recurrently mentioned frequency bands thought to reflect cognitive and memory processes in adults, children and to a lesser extent infants are alpha (8-13 Hz in adults, typically 6-9 Hz in infants) and theta (4-7 Hz in adults, 3-6 Hz in infants) band power (Saby & Marshall, 2012). Klimesch (1999) postulates that alpha desynchronization in adults is associated with semantic long-term memory performance and theta synchronisation with the encoding of new information, which is in close contact with long-term memory. Alpha and theta bands are related to one another, however in a way that they respond in opposing directions. In other words when theta synchronizes with increase in task demand, alpha

desynchronizes (Klimesch, 1999) with the probability that these two oscillations jointly contribute to successful encoding (Mölle, Marshall, Fehm, & Born, 2002). The relatively wide alpha frequency range can be further sub divided into distinct alpha sub-bands all putatively representing different cognitive processes. Alpha desynchronization of the "lower" band has previously been linked to joint attention. In adults, it has been shown to be particularly sensitive to attentional mechanisms with the function to focus on more relevant information in the information stream (Ward, 2003). In a dual-EEG study Lachat, Hugueville, Lemarechal, Conty, and George (2012) simultaneously recorded the brain activity of two adult participants interacting with each other. In this live joint attention paradigm, participants were instructed to draw their attention to either the same or an opposite object (colored LED's). The experimental conditions were performed in separate blocks. Desynchronization of signal power in the alpha range was found when both participants focused on the same object subsequent to shared eye contact when compared to the participants looking at different objects (Lachat et al., 2012). Hoehl and colleagues (2014) reanalyzed existing ERP data from Striano et al. (2006), looking at oscillatory brain activity with the aim to examine whether an increase in brain activity to objects would occur after the episode of joint attention. Their results indicated that desynchronization of alpha band activity over fronto-centro-parietal channels (in the lower band 5-7 Hz) is indeed present in 9-month-olds when the adult had previously engaged in eye contact prior to turning to the object presentation. This leads to the suggestion that eye contact induces increased brain activation in response to an object (Hoehl et al., 2014). It may although be questionable if alpha suppression during the presentation was actually related to information processing about the object as a consequence of joint attention or was due to the arousal caused by the eye contact itself. More recently, Michel and colleagues (2015) also

found alpha band desynchronization to be more supressed in 4- and 9-month olds to faces with object directed eye gaze when compared to averted eye gaze. The authors postulated that this frequency band might reflect early social object learning mechanisms.

Desynchronization of activation in the "upper" alpha range in adults has further been associated with accessing a semantic knowledge system, which could play a role in learning (Klimesch, 2012). Semantic orientation enables individuals to selectively access stored information such as sensory information and "higher order information", including language, mathematics, and geography (Klimesch, 2012) with its amplitude being larger for semantically integrated information. In a study by Klimesch and colleagues (1997) adults were asked to judge feature-concept pairs such as claws-eagle or wings-banana according to semantic congruency. When further looking at individual performance the authors found "upper" alpha desynchronization to be significantly larger for good semantic memory performers when contrasted to bad performers. Therefore, alpha-band desynchronization during a joint live interaction involving eye contact as found by Lachat et al. (2012) and Hoehl et al. (2014) may relate to the activation of the semantic knowledge system as postulated by Klimesch (2012) and Michel et al. (2015). Overall, we can conclude that frequency analysis techniques are indeed a useful tool to help us understand the underlying neural mechanisms of social information processing, with enhanced alpha desynchronization indicating that attention is being allocated to relevant aspects of the social context. This enables object learning in these situations (Michel et al., 2015). Given Hoehl et al. (2014) and Michel et al. (2015) both looked at a smaller alpha band namely 5-7 Hz we asked if perhaps there are indeed distinct sub-bands in the alpha range in infancy as reported in adults by Klimesch (1999). To date no publications have explicitly examined this issue.

Theta band synchronisation is thought to be involved in executive and voluntary control of attention (Orekhova et al., 1999) and has further been related to the encoding of new information and successful recall. Klimesch and colleagues (1996) found that when there was an increase in theta power during the encoding of a word, those words were more likely to be remembered in a subsequent recall task. The study was designed to foster incidental learning to control for any additional task-related increase in theta power. Consequently, participants were unaware that following their initial participation in a simple semantic category judgement task ("yes" for living e.g. birds; "no" for non-living e.g. vehicles) while EEG was recorded continuously, they will be asked to recall the 96 words previously seen in the judgement task. The authors concluded that remembered words must refer to the established memory. In infancy theta synchronization was observed to be greater over frontal channels during a peek-a-boo game when contrasted with quiet visual attention such as watching soap bubbles (Orekhova et al., 1999) and is further increased in amplitude during reaching and handling of objects (Fugati et al., 1998). These observations have led to the proposal that infant theta activity may be linked to learning (Saby & Marshall, 2012). In a recent a study Begus, Gliga and Southgate (2016) found an increase in theta activation in situations in which 11-month-old infants anticipated the reception of information about an object. Begus and colleagues (2015) have further demonstrated that there is a relationship between theta and learning in infants, with more pronounced theta activation during object observation leading to a more superior object recognition over frontal sites at 3-5 Hz.

It is currently unknown if "upper" alpha-band desynchronization could represent a potential and *complementary* neurophysiological marker for the N400 ERP component in early development. We further aimed to investigate theta band

synchronisation as a possible reflection of learning and explore how theta as a representative neural marker for "encoding" may relate to the formation of new information during infancy (Klimesch et al., 1996). Lastly, we aimed to investigate possible attentional differences via the Nc component as a function of consolidation. For this purpose, we designed a combined ERO-ERP within-subject design study. On visit one an experimenter first introduced and labelled two novel objects and additionally used referential communication (infant directed speech, eye gaze and naming the learner by his or her name) to facilitate learning in a 6-minute live interaction (learning phase). We investigated infants' memory performance immediately after and with a 24-hour delay by presenting the novel nouns from the learning phase via loudspeakers followed by one of the two objects seen in the learning phase. The word-object pairs were either congruent or incongruent. We postulated that mismatches (incongruent condition) would be detected for word-object pairs semantically as indexed by the N400 and would further be evident by greater allocation of attention as indexed by the Nc component. Word-object matches (congruent condition) would on the other hand be processed semantically via alpha desynchronization. If retrieving knowledge from the knowledge system was only observed in the delayed retention phase as a function of offline consolidation, then we would hypothesize that theta synchronisation is going to be evident in the immediate retention phase as a potential indicator of on-going encoding processes for new information. For the Nc component we expected that only fully encoded nouns would lead to a larger amplitude effect for incongruous word-object pairs. We further explored potential relations between the neural correlates. In other words, we explored the possibility whether infants who may display higher theta synchronisation on day 1 are the same infants likely to display more alpha desynchronization to semantically congruent stimuli and/or the N400 ERP

component to semantically incongruous stimuli on day 2. Thus, we investigated if within this highly social live learning paradigm (Lachat et al., 2012; Hoehl et al., 2014), upper alpha activity in early development leads to the activation of the semantic knowledge system as postulated by Klimesch (2012). Evidence of alpha band suppression would suggest that this specific manifestation of alpha is a signature of constructing the knowledge system at 10- to 11-months.

3.2 Methods

3.2.1 Participants

The final sample consists of twenty monolingual English-speaking 10-11-month-olds (mean age: 328.7 days, range: 305-348 days; 10 female, 10 male) recruited in and around Lancaster, UK. Each participant's legal guardian was given a detailed description of the study in form of a participant information sheet prior to participation. Parents'/ legal guardians' formal consent was attained prior to testing and each family received travel reimbursements for both visits, if attended, and a book as a gift for participation. All infants were of normal birth weight (> 2500g) and were born full term (37-42 weeks). Thirteen additional infants were tested but excluded due to inattentiveness to the experimental stimuli accompanied by strong movement artefacts resulting in fewer than the required minimum of artefact-free trials for either one or both of the testing sessions.

3.2.2 Stimuli

Two novel objects (a rounded one and a tipped one) were selected for this study (see Figure 3.1). The dimensional properties of the objects were held similar (rounded object = 11.4 x 3.2 x 11.4 cm, tipped object = 12.1 x 4.4 x 10.2 cm) and stimuli were presented

in real size during the experimental EEG/ERP session subsequent to the learning phase. The purpose of using only two objects was to simplify learning for infants, requiring them to map only two word-object pairs. Both novel objects were labelled with pseudo-words (YOK and BLAP), which was counterbalanced across participants. The two objects were always labelled consistently for each participant.

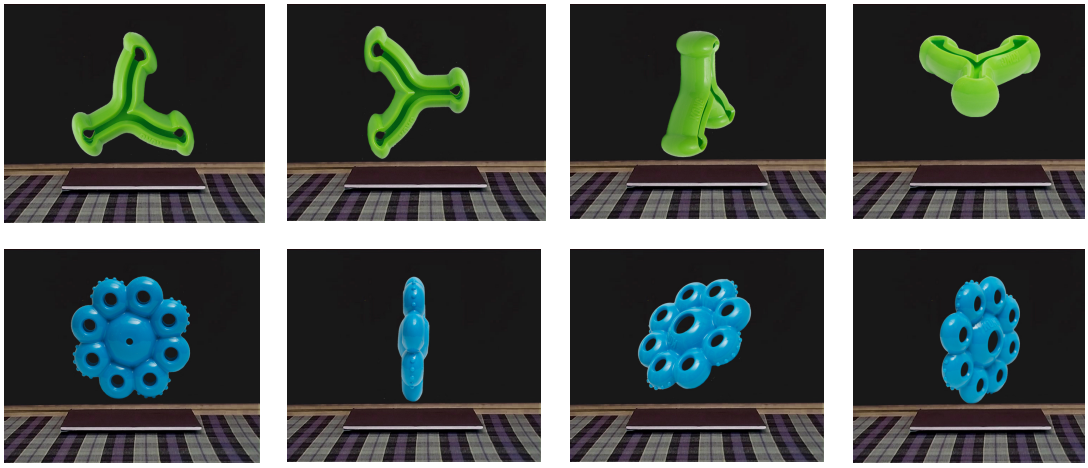

Figure 3.1: *Novel objects presented to the infants in four different angles to keep the experimental session interesting and the infant attentive. Due to infant's exploring the two objects during learning, the aim to present different angles was to facilitate recognition.*

3.2.3 Procedure

3.2.3a Learning phase

For the learning phase infants were first seated at a table on the parent's lap directly facing the experimenter in the experimental area. EEG was recorded continuously with a 128-channel Geodesic Sensor Net (HCGSN 130; EGI, Eugene, OR) for approximately 6 minutes. The behaviour of the infant and experimenter was video recorded throughout the learning phase from two camera angles for offline coding of looking behaviour to the experimenter and the object and to control for the amount of motor activity in each condition, and to measure the utterances of the word labels (Figure 3.2).

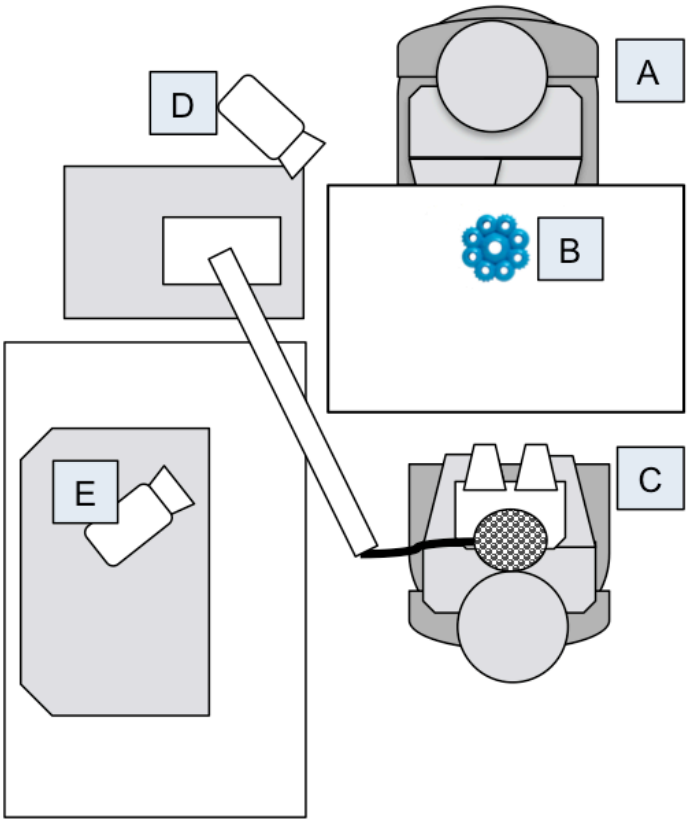

Figure 3.2: *Paradigm depicting the initial learning phase with (A) experimenter, (B) novel object, (C) the infant participant sitting on the caregiver's lap, and (D, E) video cameras.*

Both novel objects were presented and labelled in a live way and were accompanied by eye contact between the experimenter and participating infant. Objects and labels were counterbalanced across the sample. Infants first came into contact with the experimenter with the start of the learning phase. Each of the two novel objects and their novel labels were presented three times for a one-minute duration in an ABABAB fashion (total duration = 6 minutes). This procedure was chosen in order to extend each infants attention to the objects by keeping the task interesting throughout the learning phase. To account for potential order effects the order in which the objects were introduced was also counterbalanced. The experimenter tried to create a learning experience that was as naturalistic as possible. At the same time the experimenter's task was to label the novel

object as often as possible (average utterances per 1 minute block were 17.3 (*SD* = 1.3) with an average of 103.8 utterances across the 6 minutes (*SD* = 8.04). Each of the 6 blocks started by revealing one object at a time from under the table, holding up the object and saying, *"Look what I have here!"* and labelling the novel object by saying, *"Look! It's a YOK/BLAP!"*. The experimenter would first come into eye contact with the infant before the introduction sentence. Then the experimenter would bring the object closer to the infant, enabling the infant to reach out for the object to explore and make use of other social cues such as calling the infant by their name (*M* = 17.1, *SD* = 4.53 across the 6 minutes). After one minute passed (signalled to the experimenter via a short beep) the first object was then taken away and placed out of the infant's sight. The experimenter had a 10 s window to retrieve the object from the infant if needed. The start of the introduction of the second object was again signalled via a short beep. This above procedure was repeated four more times until each object was presented for 3 minutes in total.

3.2.3b Retention phase

On day 1, infants and their caregivers turned 45 degrees to face the CRT-19-inch experimental monitor. On day 2 (24-hours later), both infant and caregiver were seated facing the same monitor directly at a viewing distance of 90 cm. (Figure 3.3).

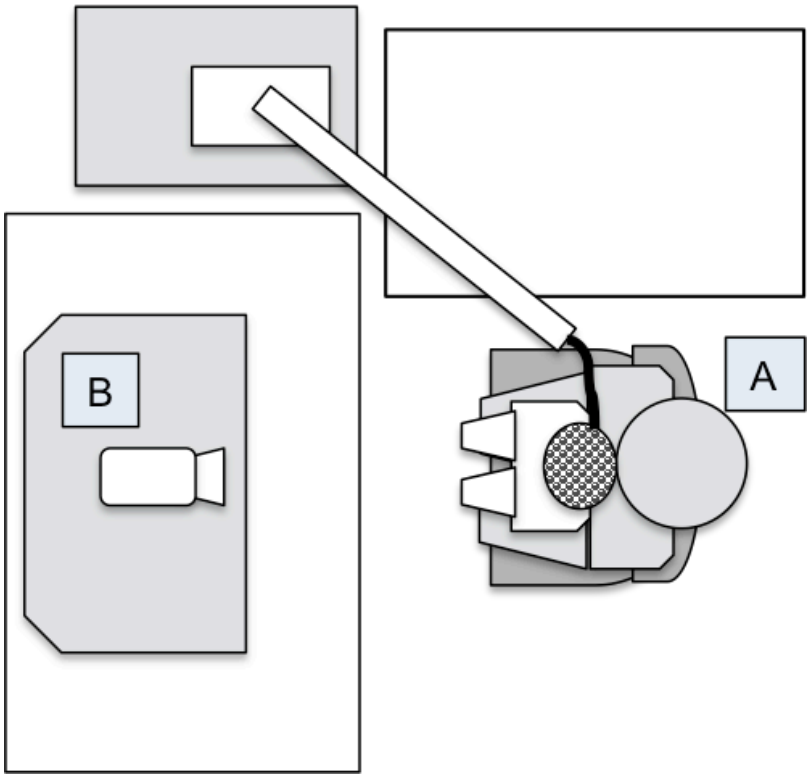

Figure 3.3: *Experimental paradigm shortly after learning and 24-hours later. The infant participant sat on the caregiver's lap (A) in front of an experimental monitor with video camera (B).*

Infants were presented with a 3-D scene in which an occluder was visible at all times throughout the testing session, positioned on a patterned surface on a black background (sequence presented in a similar fashion to Parise et al., 2008). Each trial started with an introductory phrase presented by a female neutral voice, who was not the experimenter, by saying *"This is a"* (duration: 750 ms). After this introduction and following a 50 ms pause, the prime "Yok" or "Blap" (duration: 905 ms each) was introduced by the same female voice via loudspeaker. An attention getter (black and white moving spiral) appeared on the top half of the occluder for a random duration of 400-600 ms before the occluder slowly started to drop down (350 ms) until one of the two objects from the learning phase became and remained fully visible for 1000 ms. The time lock was set to the first appearance of the object (105 ms after the occluder started to move down).

Finally, the occluder moved upwards (350 ms) until the object was fully occluded again. The next sequence started with a 400 ms inter stimulus interval (Figure 3.4). In half of the presented trials, the object matched the preceding auditory noun (congruent condition) in the other half, it did not (incongruent condition). In the event that infants became inattentive, the experimenter could either introduce a large moving spiral via key press or could pause the experiment. The experimental session lasted until the infants' attention could no longer be prolonged.

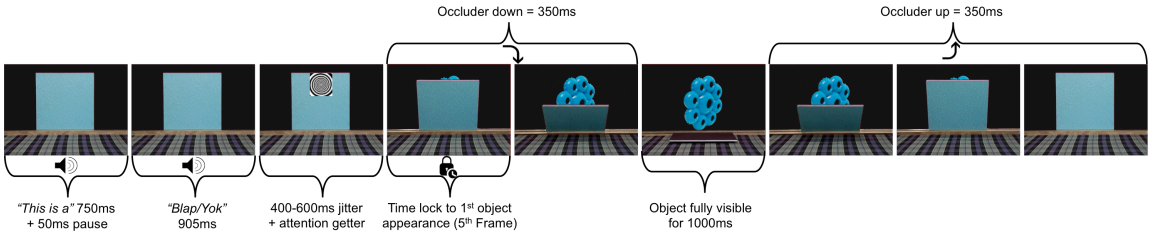

Figure 3.4: *Experimental sequence depicting the 3-D occlusion scene. The sequence always started with an introductory phrase, followed by the word prime and ended with the object probe. Word-object pairs were either congruent or incongruent. All analyses were time locked to the first occurrence of the object.*

3.2.4 EEG recording and analysis

EEG was recorded via 128-channel Geodesic Sensor Nets and amplified with an EGI Net Amps 400 amplifier with a sampling rate of 500 Hz (HCGSN 130, EGI, Eugene, OR, USA). EEG was referenced online to the vertex electrode (Cz), and an analogue bandpass filter was applied (0.1 to 100 Hz) before storing data for subsequent off-line analysis. Initial ERP data processing was performed using Net Station 4.5.4 Waveform Tools and first included digital filtering (0.3-30 Hz bandpass). Data was then segmented with a 200 ms baseline and a 1000 ms period post stimulus presentation and an artefact detection tool was applied which marked affected channels as bad whenever amplitude exceeded the set threshold of ± 200 microvolts for the entire segment, with a moving

average of 80 ms. Trials including more than 10 channels with excessive noise levels were rejected. Noisy channels from the remaining trials, which comprised fewer than 10 bad channels, were interpolated by using the average of surrounding electrodes. Remaining artifact-free trials were re-referenced to the average electrode and baseline corrected before creating individual averages for all participants. Off-line inspection of the video-recordings ensured that only trials in which infants paid attention to the images were further processed. The remaining trials were manually and visually inspected for artifacts caused by eye or body movements using normative procedures for infant ERP studies (see Hoehl & Wahl, 2012 for further information on infant ERP methods). The grand average was produced from those 20 individuals who contributed a minimum of 4 artifact-free trials per condition for each of the two testing sessions. On day 1, infants contributed a mean total of 11.35 artifact-free trials for the congruent (range 4-25) and a mean of 11.7 (range 4-31) for the incongruent condition. On day 2, the same infants have contributed with an average of 11.1 artifact-free trials (range 4-32) for the congruent and 11.8 artifact-free trials (range 4-27) for the incongruent condition. Based on findings from Stets and Reid (2011) as well as Kaduk, Elsner, and Reid (2013) and Kaduk et al. (2016), reliable and interpretable data have previously been shown to be present within infant EEG with as few as 3 to 7 trials per condition using standard visual ERP paradigms justifying the low threshold of inclusion criteria used in the subsequently outlined analysis.

For the ERO data analysis the same artifact-free trials from the ERP analysis were used, however this time the EEG signal was segmented into epochs of -1200 ms to 2000 ms around the onset of the stimulus. EEG data was post processed using custom-made scripts in "WTools" (available on request) and EEGLab (9.0.5.6b). The time-frequency analyses were conducted performing a continuous wavelet transformation.

Complex Morlet wavelets were computed at 1 Hz frequency intervals for the frequency range 2-15 Hz. Total spectral activity was calculated performing convolutions with the wavelets on all channels. The absolute value of the results was computed and served as the dependent variable. The transformed epochs were averaged for each condition (see Csibra et al., 2000). Finally, 1000 ms at the beginning and the end of each segment were removed to avoid distortion due to the transformation. Baseline correction was performed at each frequency by subtracting the mean activity of 200 ms before stimulus onset from the signal. The grand average was calculated for both conditions and time points separately. It is important to emphasize that unlike the use of wider frequency bands, narrower frequency bands reduce the possibility that frequency specific effects go undetected. Consequently, broader band analyses need to be interpreted with caution. Future research should steer away from fixed band analyses and implement a narrower frequency bands approach. In addition, the use of individual frequency bands versus fixed bands used for the entire sample is also currently not standardised in the field. The majority of studies use broad, instead of narrow bands, which are not adjusted individually (Klimesch, 1999). Due to the lack of standardisation, we have adopted our analysis based on Hoehl et al. (2014) and Michel et al. (2015) who have implemented narrower frequency bands which were in addition, not adjusted individually for each participating infant. Such an approach may help to facilitate standardisation in the field of infancy research.

3.3 Results

3.3.1 N400 ERP Component

For the analysis of the N400, a time window around 650 to 800 ms after stimulus onset in parietal sites was chosen, based on prior literature (e.g., Kaduk et al., 2016; Reid et

al., 2009) and visual inspection of the individual and grand averages. Twelve electrodes represented the left parietal (50, 51, 52-P3, 53, 54, 58, 59, 60, 61, 65, 66, 67) and 12 electrodes represented the right parietal region (77, 78, 79, 84, 85, 86, 90, 91, 92-P4, 96, 97, 101). A typical N400 peak was observed within this window on right parietal sites on day 2 (see Figure 3.5). It is seen in the incongruent but not in the congruent condition, which is in line with the N400 morphology of prior literature (e.g. Kaduk et al., 2016, Reid et al., 2009). The data were consequently analysed using the time window analysis technique (Hoorman, Falkenstein, Schwarzenau, & Hohnsbein, 1998), which is typically performed when one condition displays a defined peak in the window of interest, whereas the other does not.

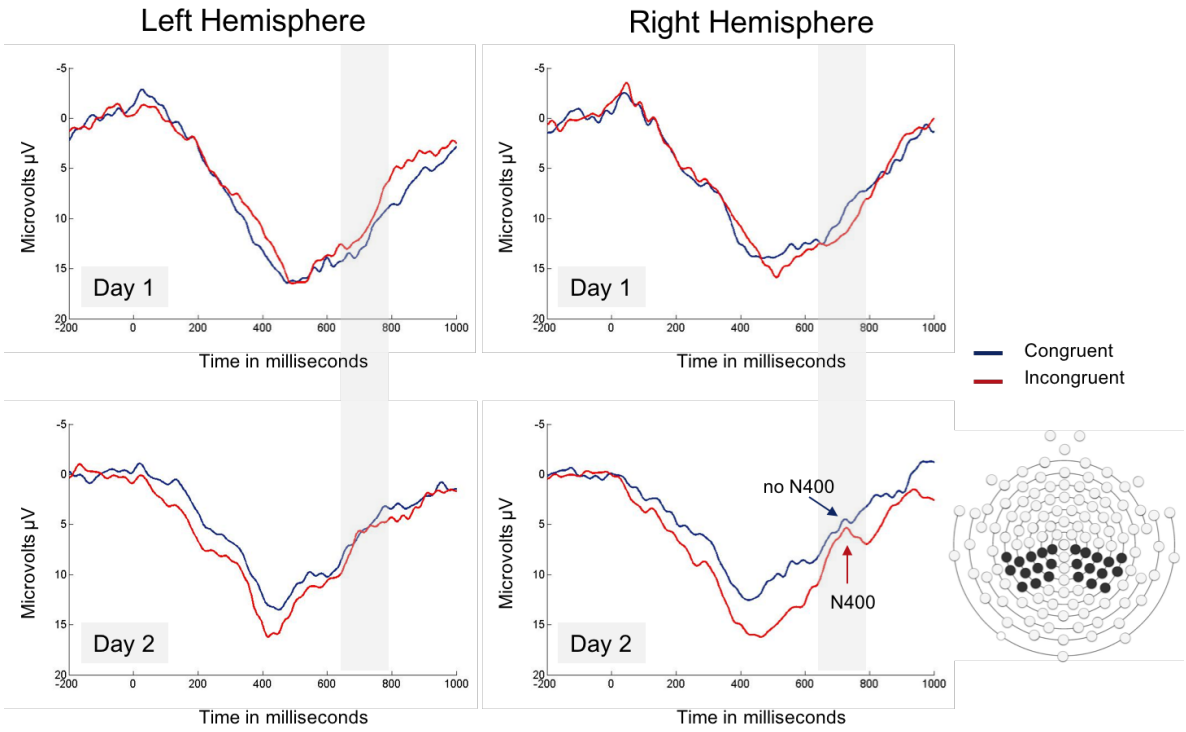

Figure 3.5: *Averaged ERPs for all infants (N = 20) from 12 channels within the left parietal region for the N400 (50, 51, 52-P3, 53, 54, 58, 59, 60, 61, 65, 66, 67) – top and bottom left and right parietal region (77, 78, 79, 84, 85, 86, 90, 91, 92-P4, 96, 97, 101) top and bottom right for day 1 (top) and 2 (bottom). Bottom right averaged channel group depicts the observed N400 for the incongruent condition only, within the time window of interest highlighted in grey.*

Therefore, time is included as an additional within-participant factor in the repeated measures ANOVA. A significant interaction between condition and time indicates validity for the hypothesis that the N400 component is evident only for the incongruent condition as a function of variance over time. A significant interaction between condition and time needs to occur to prove validity for the hypothesis that the N400 component is only evident within one condition as a function of variance over time. In order to examine day as a factor in addition to the two-way condition time effect, we would encounter statistical problems. A three-way interaction would be very difficult to establish when there is an unequal amount of variance across the two days, which is what is predicted and shown from this type of analysis. The variance on day one would need to be considerably more homogenous for a three-way interaction to emerge. In cases like this, when looking directly at variability in developmental data it is strongly recommended to look at data separately (Thomas et al., 2009). The repeated measures ANOVA, featured the two factors condition (congruent or incongruent) and time (16 samples at 1 sample per 9 ms). Due to a narrower time window 650 - 800 ms as opposed to 600 - 800 ms, and to make up for the 50 ms difference, one sample was subtracted from the total of the 17 reported in previous action N400 literature (e.g., Reid et al., 2009). Consequently, we performed a 2 x 16 repeated measures ANOVA with the between subject factors Condition (Congruent, Incongruent), and Time (16 samples) for the two days and scalp locations (left and right) separately. Results revealed a significant Condition x Time interaction only for the right hemisphere on day 2, $F(3.4, 64.4) = 3.149$, $p = .026$, partial $\eta^2 = .142$. No significant interactions were evident for the left hemisphere at the same time point (day 2), $F(3.2, 60.4) = 1.429$, $p = .242$, partial $\eta^2 = .070$ nor for day 1 of the left, $F(3, 57.9) = .603$, $p = .618$, partial $\eta^2 = .031$ or right hemisphere $F(3, 57.8) = 1.441$, $p = .240$, partial $\eta^2 = .070$.

3.3.2 Nc Component

Visual inspection of the grand average indicated frontal Nc effects more predominantly over right frontal sites 400-600 ms after stimulus onset on day 2 only. We computed the mean amplitudes for the following 13 left frontal (18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 33, 34) and 13 right frontal (2, 3, 4, 8, 9, 10, 14, 116, 117, 118, 122, 123, 124) channels and performed a 2 (condition) x 2 (hemisphere) x 2 (day) repeated measures ANOVA. The results indicated a significant Condition x Hemisphere interaction, $F(1,19) = 4.419$, $p = .049$, partial $\eta^2 = .189$. No other main effects or interactions reached significance. To explore the interaction, we therefore performed a 2 (Condition) x 2 (Hemisphere) repeated measures ANOVA for the two days separately. For the analysis on day 2, results indicated a significant main effect for condition, $F(1,19) = 4.713$, $p = .043$, partial $\eta^2 = .199$. No other significant main effects or interactions on day 2 or day 1 reached significance. Due to the lack of evidence for hemispheric differences on day 2, we collapsed the two hemispheres together. This shows a larger amplitude effect for incongruent (M = -16.15 μV, SD = 7.59 μV) compared to congruent (M = -12.9 μV, SE = 8.57 μV) word-object pairs (Figure 3.6).

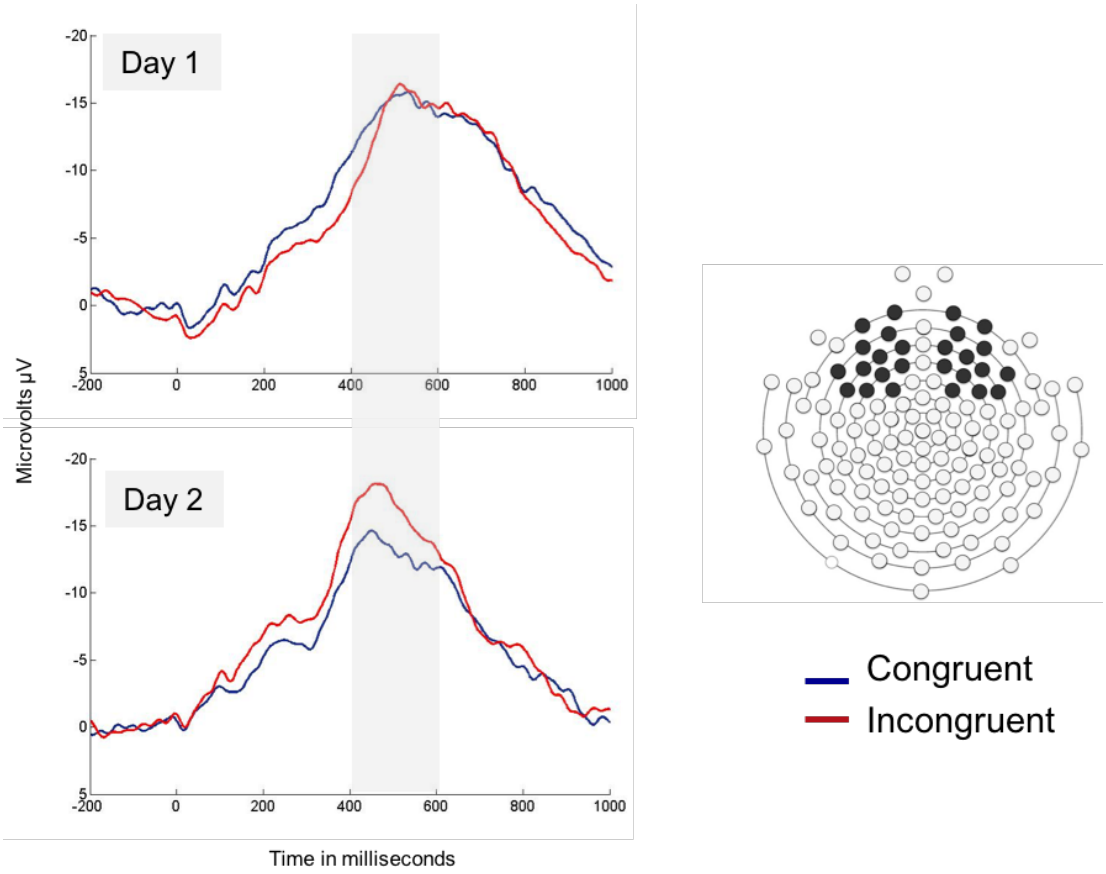

Figure 3.6: *Averaged ERPs for all infants (N = 20) from 26 channels collapsed across the two hemispheres with the analysed area highlighted in grey. Averaged ERP for day one is displayed on the top left and day 2 on the bottom left.*

3.3.3 Theta Band Power

For the analysis of the theta band, nine channels representing frontal sites (4, 9, 10, 11-Fz, 15, 16, 18, 19 and 22), and a time window around 400 to 1000 ms after stimulus onset and a frequency band of 3-5 Hz was chosen, based on prior literature (Begus et al., 2015) and visual inspection of the individual and grand averages (Figure 3.7).

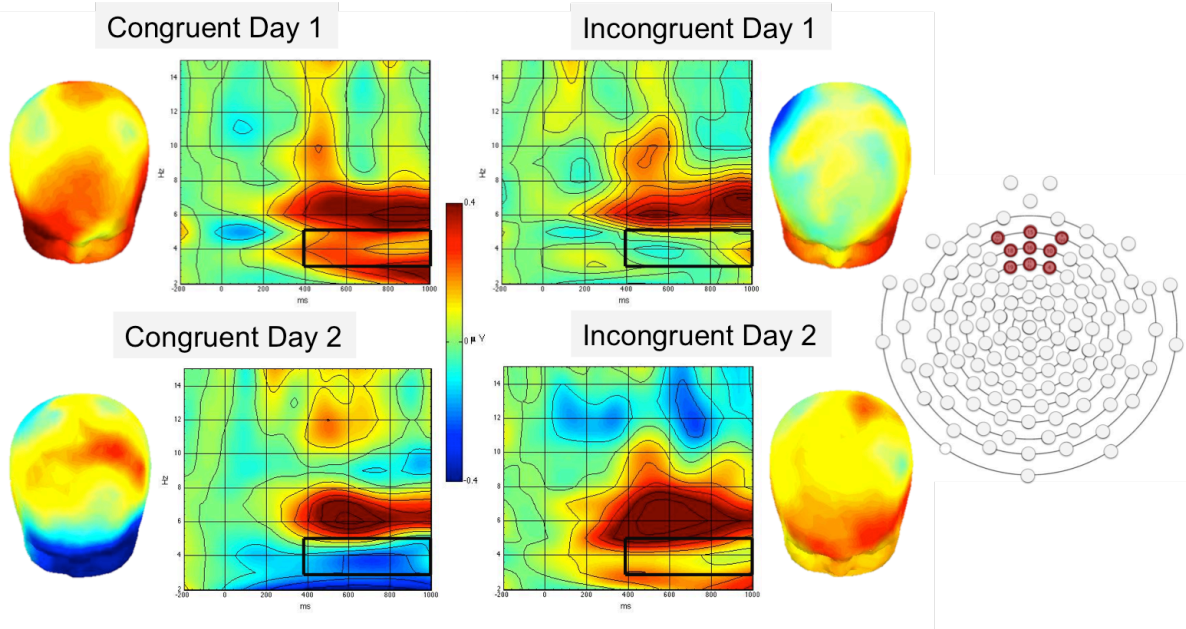

Figure 3.7**:** *Theta power presented separately for day 1 and day 2 for the congruent (left) and incongruent (right) conditions for 9 frontal channels (4, 9, 10, 11, 15, 16, 18, 19, 22). A black frame ranging from 3-5 Hz from 400-1000 ms post stimulus presentation highlights the observed and analysed theta band. 3D Scalp Maps (superior view) were computed for each condition and day representing the mean of the time window and frequency range of interest. The analysed channels are highlighted in red.*

First we performed a 2 (condition) x 2 (day) repeated measures ANOVA, which revealed a significant interaction between condition x day, $F(1, 19) = 13.05$, $p = .002$, partial $\eta^2 = .407$. Follow-up paired sample *t*-tests comparing the two conditions for both days separately indicated no significant difference between the congruent and incongruent condition on day 1, $t(19) = 1.754$, $p = .095$, $d = .239$ ($M_{congruent} = .236$, $SD = .68$, $M_{incongruent} = .023$, $SD = .61$) or day 2, $t(19) = -1.89$, $p = .073$, $d = .583$ ($M_{congruent} = -.101$, $SD = .47$, $M_{incongruent} = .21$, $SD = .59$). Comparisons for each condition and day with baseline using one-sample *t*-tests yielded no significant results.

3.3.4 Alpha Band Power

Present literature reports alpha occurring in posterior-occipital sites with the frequency typically ranging from 6-9 Hz (Cuevas et al., 2014) with upper boundaries of 10 Hz also being reported in infant literature (Stroganova & Orekhova, 2007). Based on visual inspection of the differences between conditions, data indicated activation in the upper band of alpha at around 9-10 Hz in 8 parietal-occipital channels (67, 70, 71, 72, 75-Oz, 76, 77, 83) at 700-1000 ms after stimulus occurrence (Figure 3.8).

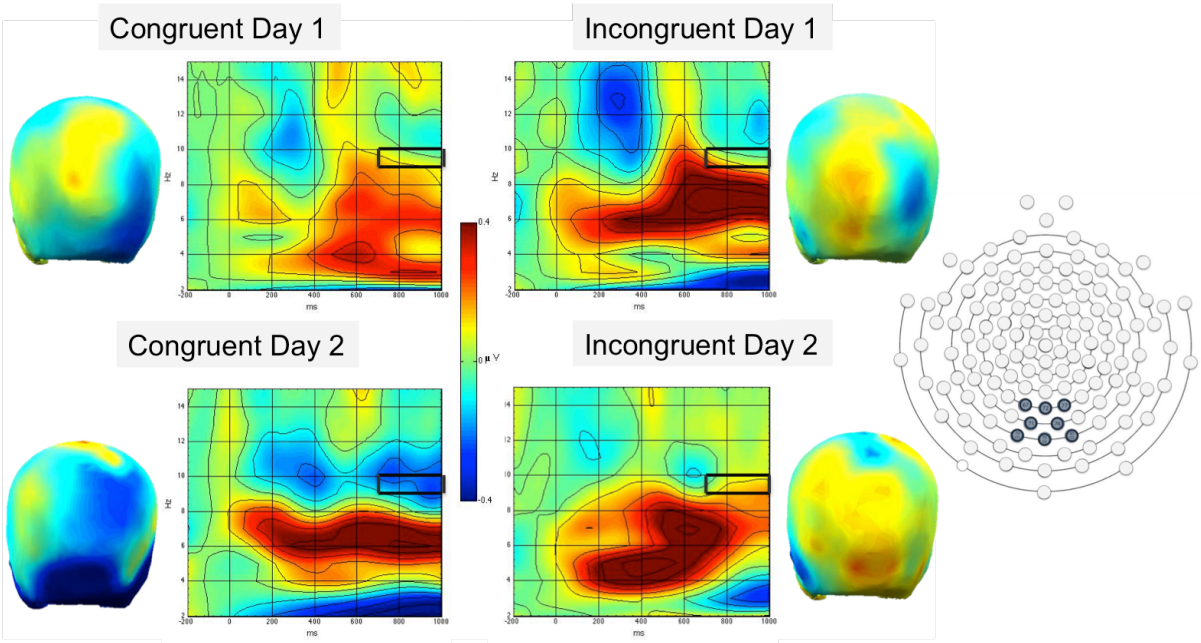

Figure 3.8**:** *Alpha power presented separately for day 1 (top) and day 2 (bottom) for the congruent (left) and incongruent (right) conditions averaged over posterior channels. A black frame ranging from 9-10 Hz from 700–1000 ms post stimulus presentation highlights the observed and analysed alpha band. 3D Scalp Maps (posterior view) were computed for each condition and day representing the mean of the time window and frequency range of interest. The analysed channels are highlighted in blue.*

A 2 (condition) x 2 (day) repeated measures ANOVA revealed a significant main effect for condition, $F(1, 19) = 7.940$, $p = .011$, partial $\eta^2 = .295$. Neither a main effect for day,

nor a significant interaction between condition and day was found. Follow up paired sample *t*-tests showed that the main effect for condition was driven by significant differences between the conditions on day 2 only, $t(19) = -2.902$, $p = .009$, $d = .383$ with the congruent condition displaying more negativity ($M_{congruent} = -.205$, $SD = .345$) when compared to the incongruent condition ($M_{incongruent} = .088$, $SD = .26$). One-sample *t*-tests (with the test value = 0) revealed that only the congruent condition on day 2 reached significance from baseline, $t(19) = -2.655$, $p = .016$.

3.3.5 Correlation Analysis

In order to correlate the N400 with the remaining variables of interest and significant outcome (Nc & alpha for day 2) it was first necessary to find a representative value for the observed variation over time for one but not the other condition from the time applied window analysis technique (Hoorman et al., 1998). For this purpose, we computed a measure of deflection from the response curve (deviations) by summing the vertical difference between the curve and a straight line between the two end-points (either side of the identified time window: 650 ms - 800 ms post stimulus onset) of the curve at each sampled time (75 samples in total) between those two points (Figure 3.9). This resulted in a set of deviation values for which positive values characterized a larger deflection in the negative direction (N400) and values close to zero or of negative value represent an absence of the N400 amplitude.

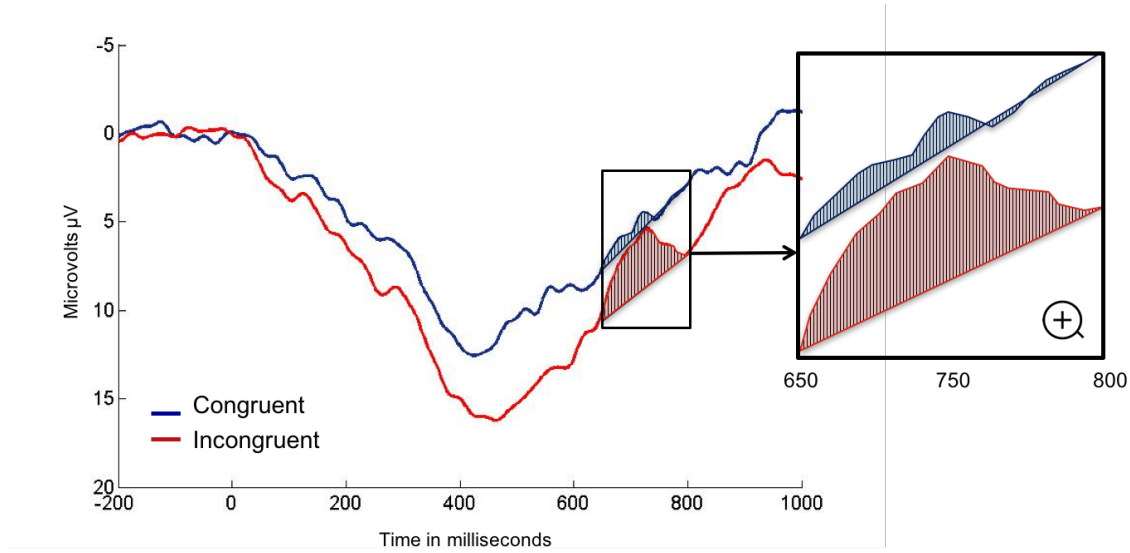

Figure 3.9: *Averaged ERP's for (N = 20) subjects over right parietal channels on day 2. The framed area depicts the analysed time window of interest displaying a more pronounced N400 for the incongruent condition (red line). A line was drawn at the two points either side of the time window in which the ERP crossed the border for each individual, condition and day. A measure of deflection was computed from the response curve by summing the vertical difference between the curve and a straight line between the two end-points of the curve at each of the 75-recorded samples.*

In addition to comparing differences in power between conditions, all conditions should be also measured in respect to the baseline against zero. A lack of significance from baseline can cause problems for the interpretability of the results. One potential possibility could be that one of the two conditions could show an increase in power (synchronization) and the other a decrease in power (desynchronization). As a consequence, the significant difference between conditions could be driven by the opposite power difference and not the desynchronization of one condition (see Cuevas et al., 2014). As the present study failed to show any significant difference from baseline for any of the theta related variables or alpha on day 1, we are not reporting any additional correlations in the subsequent correlation analyses.

Subsequently we calculated the difference between conditions for our target

variables. For the N400 we specifically focused on day 2 over the right hemisphere, as this was where the evidence for the N400 occurs between conditions in the present dataset ($N400_{incongruent}$ - $N400_{congruent}$). The second ERP variable of interest, which was entered in the subsequent correlation analysis, was the Nc component, which elicited differences in attentional processes again on day 2 over frontal channels. The difference score for the Nc resulted from the collapsed frontal channels on day 2 ($Nc_{incongruent}$ - $Nc_{congruent}$). The only value entering the correlation for the ERO's was the upper alpha difference in power between conditions for day 2 ($alpha_{congruent}$ - $alpha_{incongruent}$). Given our very strong a priori predictions we performed one-tailed bivariate (Pearson's) correlations to all remaining comparisons. Results revealed a significant correlation between N400 and alpha, r = -0.454, p = .022, (Figure 3.10). No other correlations reached significance.

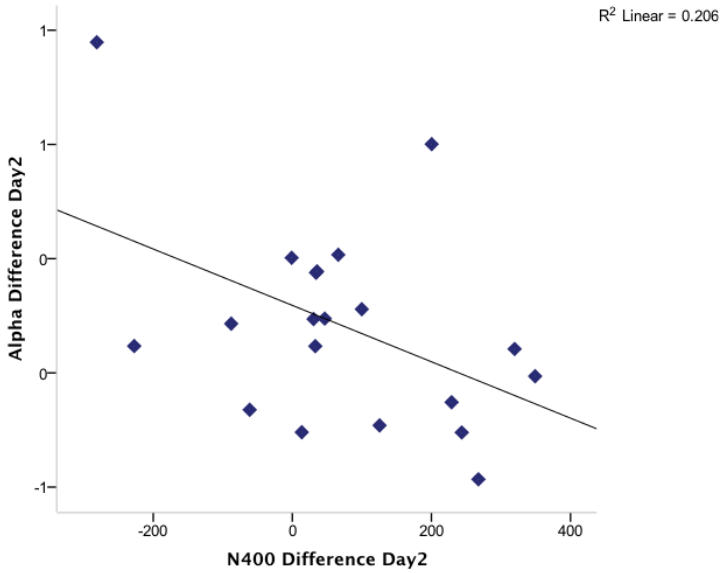

Figure 3.10: *Correlation scatter plot depicting the two variables N400 deviation value difference (Incongruent - Congruent) on the x-axis and alpha band power difference (Congruent - Incongruent) on the y-axis. Larger relative values for the N400 represent the infant's ability to detect word-object mismatches in the delayed retention phase. The same infants further displayed more negativity in the alpha range in the congruent relative to the incongruent condition.*

3.4 Discussion

The objectives of the present study were to investigate complementary neurophysiological modulators of semantic processing in early development. More specifically we aimed to target "upper" alpha-band suppression using ERO's and relate this frequency band, thought to reflect activation of the semantic knowledge system in adults (Klimesch, 2012), to the N400 ERP component during infancy. The N400 has consistently been reported to represent the processing of meaning from infancy to adulthood in both the action and language domain (see Amoruso et al., 2013; Kutas & Federmeier, 2011 for a review). Despite numerous attempts however, and most likely due to differences in paradigms, the evidence of the N400 from developmental literature is somewhat inconsistent. Thus, the need for further investigation into the N400 phenomenon was palpable. The social aspect of the live presentation within the learning phase was chosen to demonstrate how upper alpha desynchronization as well as the N400 ERP effects later information retrieval. We aimed to test if a putative mechanism of semantic rule use is applied for making inferences about newly acquired words, and their corresponding verbal referents for both of these measures. We focused on the language N400 ERP component, a component that is equally evident in action-based violation of expectancy paradigms and aimed to contribute to our understanding on these two N400 measures (action and language N400) to explore if they possess a common source as suggested by Amoruso et al. (2013).

We hypothesized that we would find alpha band desynchronization of the upper band (9-10 Hz) in early development similar to that of adults (Klimesch, 2012) leading to the suggestion that the specific manifestation of alpha may well represent a signature of constructing the knowledge system also in infancy. We looked at possible relations between these two correlates to help answer the question if both neural markers index

the same cognitive mechanism, namely the processing of meaning via the use of a semantic expectancy task. Further, we considered theta band synchronisation (3-5 Hz) to be a possible reflection of learning and explored theta as a representative neural marker for active and selective preparation of information encoding as for instance postulated by Begus and colleagues (2016). In addition, our investigations extended to infants' attentional processes. In other words, infants' ability to differentiate between newly learned word-objects pairs using the Nc ERP component as a useful tool. Lastly, we aimed to contribute to our understanding of the trajectory of semantic formation. For this purpose, we introduced two retention time points (immediately after learning and with a 24-hour delay). We hypothesized that semantic formation evolves as a function of consolidation processes (Friedrich et al., 2014), which stands in accordance with the dual memory system literature (Henderson, et al., 2012; 2013). For this purpose, twenty 10-11-month-olds engaged in a live learning situation for 6 minutes in which they observed an experimenter introduce and label two novel objects. Infants' memory performance was tested immediately after and with a 24-hour delay. First a mental representation was induced by priming infants with one of the two nouns from the learning phase (Yok/Blap). Subsequently either the correct or incorrect object appeared from behind an occluder creating compliance or violation of expectation.

The main finding of this study comes one step closer in supporting the assumption that upper alpha frequency band over posterior sites could well relate to infants ability to retrieve semantic information from the semantic knowledge system as postulated by Klimesch (2012). In support of this notion, upper alpha power showed more suppression during the processing of congruent word-object pairs (semantic alpha). Further our finding revealed the presence of the N400 ERP component for incongruent word-object pairs as expected in posterior sites in the typical time window

and with the typical morphology (N400 in the incongruent condition only) peaking at around 700 ms post stimulus presentation (see also Reid et al., 2009, Kaduk et al., 2016 for similar results). We suggest that 10-11-month-olds were able to form semantic representations from a short 6-minute live learning situation after a one day delay.

An initial evaluation would suggest that contradicts the findings from Parise and Csibra (2012) who proposed that 9-month-olds detect the mismatch between familiar objects and a preceding label only if their own mother, but not an experimenter, introduced those objects. At this point it is important to reason about the difference in channels selected for this analysis when compared to the original study by Parise and Csibra (2012). Firstly, the study by Parise and Csibra (2012) motivated our paradigm due to their use of a less frequently used method to apply the semantic violation paradigm for the investigation of the language induced N400 ERP, namely by introducing the language prime first. We believed that this approach would reflect a more accurate and consequently more genuine access to semantic knowledge as infants in this study would first need to recruit their knowledge system via a more abstract route – by forming a mental representation induced by auditory information. Secondly, unlike the words use in the original study, words presented to the infants were not familiar and had to be learned in a 6-minute-long live interaction on the initial visit. In our paradigm, we did not introduce the words in the EEG acquisition phase in a live manner through the mother or an experimenter in a child directed fashion but instead used a neutral female voice, unknown to the infant participants, via loudspeakers. Infants in the present study were further of a slightly older age. Although the channels chosen to enter the N400 analysis differ from Parise and Csibra's (2012), nonetheless, five out of eleven channels in the right and five mirroring channels out of eleven on the left did overlap across the two studies. It is interesting to note that both studies observed

N400 effects that were dominant over the right hemisphere. This stands in accordance with the reported right bias for the N400 phenomenon, which is further parietally maximal in adult literature (Kutas & Federmeier, 2011). The present dataset thereby contributes to our general knowledge of the infant N400 with the finding that stable mental representations can indeed be formed and retrieved from semantic memory, even if the objects were not primed by the mother (Parise & Csibra, 2012) and when taking offline consolidation into account (Friedrich et al., 2014).

Furthermore, this study demonstrated that alpha desynchronization and the N400 are to some extent related to one another, meaning that those infants who displayed an N400 for incongruent word-object pairs in the delayed retention test (24 hours later) were also more likely to show greater alpha suppression for congruent word-object pairs when compared to incongruent pairs. This leads to suggest that both neural markers may well represent related cognitive mechanisms as both measures were previously linked to index semantic processing. However, due to the correlational nature of this aspect of the study, we cannot comment on any causal relationship between the two measures. The presence of alpha band suppression suggests that this specific manifestation of alpha is a signature of constructing the knowledge system at 10- to 11-months as postulated by Klimesch (2012) in adults, further suggesting the possible existence of multiple sub-bands in the alpha range. We believe that both measures are sensitive measures of semantic violation thus they are two independent measures of the same construct. Within the field of developmental cognitive neuroscience, we now have two different techniques for indexing semantic expectancy and therefore similar neural mechanisms. In order for a stronger conclusion to be made, further research needs to be pursued.

The current study further revealed an increased negativity of the mean amplitude of the Nc for incongruent word-object pairs again only 24-hours after the learning phase, indicating that this correlate not only relates to novelty or familiarity as previously thought (e.g., de Haan, 2007) but could well respond to violation of expectations of learned associations, indicating potential associative memory as reported by Nordqvist, Rudner, Johansson, Lindgren, and Heimann (2015). A more recent line of research has investigated possible links between the Nc and the N400 in which 20-month-olds had to map pseudo words to novel object referents (Borgström et al., 2017). The authors reported a predictive function of the Nc component to novel objects and the N400 incongruity effect and proposed that efficient early visual processing of objects is linked to word learning abilities. The present study failed to yield a significant relationship between the two correlates other than the simultaneous presence of both correlates on day two. One possible explanation could be that the difference reported in Borgström and colleagues (2017) related to the Nc differences between conditions for the first three presentations only. In other words, fast-mapping abilities over all artifact-free trials did not indicate any stable links with the N400 component. In the present study, we have only looked at overall Nc amplitudes due to our main focus on possible relation between the N400 and alpha power.

Consistent across all significant results we found evidence for semantic and memory processes only in the delayed retention test. This leads us to conclude that in the first instance (immediately after the learning phase) some lower level associations must have been formed. These associations, however, were not strong enough to enable 10- to 11-months old infants to semantically process information and evaluate correct from incorrect word-object pairs. With a 24-hour delay, infants have undergone additional processes related to offline consolidation, which were evident in infants'

increased memory performance (Nc, N400 & upper alpha suppression). Our findings support the notion of the existence of a dual memory system (Henderson et al., 2012; 2013) with the importance of offline consolidation processes (Brown et al., 2012) helping to reorganize recent memories for the formation of semantic knowledge (Friedrich et al., 2014). To date, much of the evidence on early word meaning acquisition is based on online encoding processes (Wojcik, 2013). Comprehension tasks that have been previously used typically required children to choose the newly learned object from an array of familiar ones. However, once children had to choose an object amongst other equally novel ones, children were unable to map the correct referent to the corresponding object (e.g., Mather & Plunkett, 2012). Comparably little is known about children's retention of word meaning representations thus it is crucial to study the properties of the semantic system beyond initial encoding (Wojcik, 2013).

The present study aimed to investigate theta band synchronisation as a possible indicator for "encoding" (e.g., Begus et al., 2015; 2016; Klimesch et al., 1996). However, we failed to find clear evidence to indicate active encoding in the immediate retention phase, as indexed by the absence of theta band synchronization on day 1. We argue that our paradigm was not carefully enough chosen for the assessment of theta synchronisation. This could well be due to the fact that encoding of information had already taken place in the 6-minute live learning phase, which makes the evidence of theta in the immediate retention phase redundant. It is further possible that encoding took place at a later stage, for instance during offline consolidation which was not considered in the given paradigm. Consequently, future research should investigate infants' theta activation during the learning phase which could further inform us how encoding and the recruitment of semantic knowledge may interact in early development. This would represent an intriguing possibility for further investigation with theta

correlating with successful learning. Under this scenario, not only the N400 could serve as a useful tool to test semantic memory but also alpha band suppression may well be fruitfully examined when investigating the ability to learn and remember what words refers to in early development.

Lastly, our findings cannot be extended to generalization and categorization abilities as our design only used two novel objects and two novel labels to simplify learning. Future research could increase complexity by introducing additional objects during the initial learning phase.

3.5 Conclusion

This study is the first study to demonstrate and to relate infant modulation of posterior upper alpha (9-10 Hz) and N400 as two representative markers for semantic processing and recruitment of the knowledge system. We are proposing a distinct time course that leads to the formation of meaning and information storage and therefore provide evidence for the notion of a dual memory system. Our findings therefore contribute to our understanding of upper alpha band suppression and its involvement in integration and retrieving knowledge from the knowledge system and support the notion of the existence of distinct sub-bands, similar to those reported in adult literature. These results strongly indicate that infants can learn semantic information rather than rely on associative learning as an underlying mechanism for social processing. In order to fully enable these processes to come to light, however, one needs to consider that consolidation of new information happens over time.

3.6 Authors Contributions

K.K., V.M.R, E.P., and S.H. conceptualized this study and K.K. developed the methodology. K.K. and K.D. were involved in data acquisition and K.K. undertook data editing of the acquired data. K.K, V.M.R., E.P., and B.T. performed the formal analysis. The original draft was written by K.K. who also provided visualizations. K.K, V.M.R., and S.H were involved in review and editing. Resources were kindly provided by V.M.R.

Prelude to Chapter 4

The link between action and language recognition is complex and bidirectional: bridging what we hear, see and do.

The recruitment of *mirror neurons* in area F5 of the macaque brain has been taken as neurophysiological evidence for a common system for action observation and action execution predominantly for goal directed hand and mouth actions. The perception of the agent or an object alone, however, does not seem to activate the same neurons, therefore suggesting that an interaction between the agent and the object is required (Gallese, Fediga, Fogassi, & Rizzolatti, 1996). Kohler and colleagues (2002) later proposed that activation of neurons in F5 also correspond to action-related sounds, postulating that these audio-visual sensitive neurons must code abstract context, namely the meaning of actions. Interestingly, action sound and purely visual representations of the same action activated the same mirror neurons. This suggests that audio-visual mirror neurons may well play a role in planning and executing as well as recognizing the actions of others (Kohler, Keysers, Umiltà, Fogassi, Gallese, & Rozzolatti, 2002). This view of action processing has motivated numerous studies in adults generating converging evidence, but also controversy. In contrast to the macaque brain, activation in suggested mirror neuron areas of the human brain does not appear to be limited to the observation of actions (e.g., Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995). Static images of implied actions, for instance, have been reported to activate areas implicated in a putative mirror neuron system (see Nishitani & Hari, 2002; Urgesi et al., 2006). As such, activation in human motor areas appears to be related to more abstract forms of event processing (e.g. gesture recognition). This is not surprising given the evidence

that describes a close link between F5 and the human Broca's area despite this area typically being considered as a speech area (Rizzolatti, Ferida, Gallese, & Fogassi, 1996). This link may represent a pivotal connection between speech and action representation (Rizzolatti & Arbib, 1998). The speech evolutionary perspective proposes a transfer of gestural hand/arm meaning to more abstract sound meaning (Rizzolatti & Craighero, 2004), which indeed sounds plausible. Locatelli and colleagues (2012) have demonstrated how training of complex manual hand actions can facilitate cognitive-linguistic skills as demonstrated by an increase in reaction times in a semantic language task following training (Locatelli, Gatti, & Tettamanti, 2012).

A wealth of adult studies inspired by the mirror neuron account has investigated how language may influence action planning, and vice versa, by examining inhibitory and facilitatory priming effects at different levels of analysis, such as single words (e.g., Gentilucci et al. 2000; Glover et al. 2004; Martin et al. 1996; Myung et al. 2006; Pulvermüller et al. 2005), and sentential meaning comprehension (e.g., Glenberg & Kaschak, 2002; Glenberg et al. 2008; Moody & Gennari, 2010; Zwaan & Taylor, 2006). In sum, results from these studies suggest that in adults the link between action and language processes is complex and bidirectional but suggests that language and meaning derive from human experience (Glenberg & Kaschak, 2002).

At first sight, action perception and language processing differ in many respects, ranging from the lack of specification in described events when compared to observed actions, through to the specific temporal dimensions of linguistic processing (e.g., events unfold over a sequence of actions, yet event descriptions do not typically unfold in the listener's mind word by word). In addition, considering for example language-induced action simulation by spoken words, it should be clear that motor activation may occur in response to the action referred to in the content of the utterance (referential

motor resonance) or/and as a result of the communicative act of speaking (communicative motor resonance, see Fischer & Zwaan, 2008 for a more in-depth discussion of this issue). However, in a recent study Moreno, De Vega and Leon (2013) demonstrated that only action language (e.g., *Now I cut the bread*) and visually presented actions (e.g., bread cutting), but not abstract language (e.g., *Now I doubt the plan*), significantly modulated mu and beta rhythms over sensorimotor regions. In this study adult participants received blocks of abstract vs. action language as well as blocks presenting manual actions carried out with familiar objects. Results serve as evidence for comparable patterns of modulation in action-related frequency ranges in response to action language and action observation.

An interesting question resulting from this finding is whether the suppression of mu oscillatory rhythms plays a functional role in learning about actions and their verbal referents. After all, humans tend to verbally describe what they are doing and what they perceive in the world around them (Bendy & Caramazza, 2011). However, research investigating the role of motor activation during language acquisition particularly in early development is scant. Whether the processing of event information is captured by mirror neurons or by an internal simulation of the observed actions, the key factor for the purposes of the subsequent experimental chapter is that activation of motor areas is not strictly dependent on visual input. This has led us to investigate the link between the ability to abstract from the observation of actions and other cognitive domains, such as language comprehension.

Therefore, the aim of the final experimental chapter of this thesis was to investigate the early link between language and motor systems. We have utilised EEG techniques as they can provide direct comparability between motor activation in adults as well as infants. Developmental neuroscience studies have previously focused on the

mu rhythm, which occurs in the alpha frequency range over central electrode sites (for a review see Marshall & Meltzoff, 2011). An observed reduction in power of the mu frequency band over the sensorimotor cortex can be observed during action observation and execution (mu desynchronization). Due to uncertainties in the literature, the interplay between action and language still requires further examination.

Chapter 4

The role of motor processes in the acquisition of new actions, verbs and sounds by 2-year-olds: an observational training study

Abstract

The ontogeny of motor resonance induced by linguistic material has not been studied to the same extent as motor resonance triggered by action observation. One way to assess motor resonance is by means of EEG, more specifically by investigating modulation of the *mu rhythm*. As EEG data acquisition with toddlers is labour intensive, literature comparing mu rhythm modulation at two time points is limited. The aim of the present study was to directly compare modulation of the mu rhythm in 24-month-olds and relate suppression of mu in response to auditory action-related verbal material in contrast to auditory non-verbal material, before and after observational training. Mu desynchronization pre-training did not differ for the two referents. Post-training, the pattern of mu activation did not parallel activation to a familiar control action verb, which was part of the toddlers' motor repertoire. However, we observed significant mu desynchronization in response to the newly acquired auditory verbal-referent. It is important to note that the training period of the present study is based on action observation rather than action execution. It therefore remains open whether active training could lead to a shift in activation. Visually induced motor resonance to the novel and trained hand actions also differed as a function of training. These findings highlight the importance of simultaneous representations of actions and verbal referents. The results of this study suggest that there are early links between motor and language systems that operate at different levels of information processing.

4.1 Introduction

Action understanding can refer to different cognitive functions such as the ability to anticipate an action, or recognize the action and its action goal (Uithol, van Rooij, Bekkering, & Haselager, 2011). At first glance, two main factors seem to drive the development of action understanding. Infants initially observe others perform actions (action observation), and with time, their active engagement with the surrounding world gradually increases (leading to action experience). The interplay between repeatedly observed actions performed by others and one's own action experience may therefore form the basis for stable associations between motor acts and sensory consequences. Over time the acquired association facilitates complex cognitive functions, such as understanding others' goals or intentions based on the motor systems of the self (see Hunnius & Bekkering, 2014 for a review). It has been argued that the roots of the ability to understand the goals of others derive from these observation and execution mechanisms (Grezès & Decety, 2001). In recent years, the traditional view of separate representational domains for perception and action has therefore been replaced by a strong interest in studying motor and higher-level processes within an all-inclusive account of cognition. Two theoretical interpretations that have emerged from these research efforts relate to the idea that cognition is grounded in both perception and action (embodied cognition), and that language comprehension relies in part on motor system activation. It can be argued that a robust way to investigate the emergence and involvement of motor systems during action observation and language processing is to have a closer look at early development.

The term *motor resonance* has repeatedly been used to describe the activation of the motor system, which is evident during the observation as well as the execution of actions. This has contributed substantial evidence for a plausible link between motor

system activation and the observer's action repertoire. One line of research proposes that motor resonance in adults may be a phenomenon only when a conspecific is detected. For example, the activation reflecting motor resonance for mouth actions of different species was limited to observing humans (and to a lesser extent monkeys) as compared to dogs (see Buccino et al., 2005). In part, this suggests that actions need to be well established in the observer's mind in order to generate motor resonance (see Buccino et al., 2005; van Elk et al., 2008).

Motor resonance has been documented in numerous studies examining action expertise in adults and to a lesser extent in infants. In adults for instance, higher levels of activation were observed in prefrontal and parietal regions of ballet and capoeira dancers that observed movement within and outside their individual expertise (e.g., Calvo-Merino et al. 2005; 2006). Similar conclusions have been drawn from studies with athletes engaged in motor imagery tasks (e.g., Fourkas et al., 2008; Orgs et al., 2008; Sacco et al., 2006; Wei & Lou et al., 2010). In a recent study, Cannon and colleagues (2014) contrasted active experience with the observation of a relatively novel tool action. In this study, adults who had acquired action experience prior to testing showed higher levels of motor activation as reflected by mu rhythm desynchronization. In another study, van Elk et al. (2008) found mu rhythm of 14-to 16-month-olds to be a function of the ability to produce the observed action. Infants that were already able to crawl showed a higher mu suppression in fronto-central areas when observing crawling actions. The reverse pattern of activation was observed in infants of the same age who were less capable or unable to crawl. This suggests that the level of mu suppression in this study depended on the ability to physically perform the observed action, or at least that there was a relation with some prior knowledge about the kinematics of this action. De Klerk and colleagues (2014) attempted to investigate how perceptual-motor

couplings in the brain develop through sensorimotor experience. One group of pre-walking 7-9-month-old infants were trained to perform stepping actions on an infant treadmill. Infants from the control group only received visual experience with the stepping actions. Infants' mu suppression was recorded before and after training to short video clips of stepping actions. In this novel pre-post training study, no effects of motor experience reflected in mu suppression were found, which led the authors to suggest that during visuomotor experience, mere associative learning is taking place (de Klerk, Johnson, Heyes, & Southgate, 2014). In sum, these studies suggest that observed actions can only be comprehended if the observer has mastered those actions themselves (Nyström, Ljunghammar, Rosander, & von Hofsten, 2011).

However, this strict view of motor processing depending on prior action experience was recently challenged by Cross and her colleagues (2011), who found motor responses to unfamiliar robotic movements in adults. Grossmann and colleagues (2012) found similar effects in their fNIRS study with 4-month olds responding to novel motion patterns; for example, robot-like motion elicited more activation when compared to human-like motions. The authors suggest that left hemisphere regions were involved in the analysis of agent motion similar to that of adults, which supports a functional specialization early in human development (Grossmann, Cross, Ticini, & Daum, 2012). In part, these findings contradict the view that actions need to be first established in the observer's mind in order to generate motor resonance (e.g., Buccino et al., 2005) and suggest that observed actions may well be understood while lying outside one's own action capacities (see Fischer & Zwaan, 2008 for a discussion of this issue).

A second theoretical framework has emerged in recent years postulating that language comprehension relies on motor system activation. This line of research

investigates the link between the ability to abstract from the observation of actions and other cognitive domains, such as language comprehension. Language-induced action simulation refers to motor resonance based on verbal input (e.g., in written form or speech perception). Work by van Elk and colleagues (2010) on language-induced motor resonance in adults highlights a correlation between motor activity and the semantic association between noun-verb pairs. The authors conclude that motor resonance during language processing is more likely to serve a predictive function during the retrieval and integration of lexical-semantic information, rather than an actual simulation of actions (Van Elk et al., 2010). Dependent on the verb (e.g., actions performed with different body parts), different cognitive processes (e.g., lexical semantic access) can be activated displaying different cortical topographies (Pulvermüller, Härtle, & Hummel, 2002). Further evidence for processing of word meaning in fronto-central areas (within the motor and premotor cortex) was provided by Hauk and colleagues (2004). When adult participants were presented with action words (e.g., *pick* or *kick*) in a passive reading task, different areas along the motor strip were activated. More specifically, either the areas for action comprehension and execution directly overlapped or they were adjacent. This finding led the authors to conclude that semantic representation is not restricted to meaning-specific brain regions, but instead represents a wider representation across the entire brain (Hauk, Johnsrude, & Pulvermüller, 2004).

Fargier and colleagues (2012) found evidence for the association account (e.g., Pulvermüller, 1999), which postulates that listening to a verbal stimulus should trigger similar patterns of brain activity as during the observation of videos depicting the same action. The observed mu suppression in this study, however, was weaker in response to verbal action stimuli over centro-parietal regions when compared to viewing the videos (displaying object-oriented hand and arm movements). Interestingly, motor activity

during the perceptual experience of the actions decreased with an increase in training (Fargier, Paulignan, Boulenger, Monaghan, Reboul, & Nazir, 2012). This observed pattern was interpreted as a form of reinstatement of brain activity. It was proposed that with a longer training time a) language induced motor activation would shift from centro-parietal to more fronto-central electrode sites and b) this activation would be induced by both verbal and visual stimuli. Overall these findings highlight the importance of the interplay between perception and action.

Although the studies discussed above demonstrate the relation between action and language processing and implicate the role of experience in motor resonance, relatively few studies have investigated motor resonance in relation to the ontogeny of language and action systems. Conceivably, in early development, learning how to execute an action goes hand in hand with learning about goals, physical realities, and other people's intentions. It has been proposed that the simultaneous representation of both action performance and auditory information during the action learning process leads to a synaptic connection between neurons in specific motor and language areas (Pulvermüller, 2005). Hereby, the literature predominantly points to the left hemisphere for language processing but the right hemisphere has also been found to comprehend language and more specifically word meaning (Federmeier, Wlotko, & Meyer, 2008), espeically for the processing of more abstract verbs when compared to simple verbs (Rüschemeyer, Brass, & Frederici, 2007).

In a study by Paulus et al. (2012) infants as young as 8 months reflected a stronger mu desynchronization when listening to an action-related sound that the infants produced themselves during a prior training phase, compared to an equally familiar non-action-related or novel sound. An interesting issue is therefore whether motor resonance in early ontogeny is relatively automatic and abstract, as has been suggested

by adult research, and whether these findings also transfer to word learning in infancy. However, in Paulus's study, the familiarized sound during training had no visual action referent. The issue of whether encountering an action-related sound alongside its visual referent would lead to motor system activation, is therefore unknown. Furthermore, the actions (shaking a rattle) in Paulus et al.'s study were not novel to the infants. Consequently, Gerson and colleagues (2015) tried to fill this gap by presenting novel actions to 10-month old infants. The participants experienced active as well as observational experience (action performed by the caregiver), both of which produced distinct sounds. After training, infants' motor activity was measured during auditory presentations of the two trained sounds and one additional novel sound. Their results revealed differences in mu power to actively and passively observed action-referents, which supports the notion of the importance of motor experience to action comprehension (Gerson, Bekkering, & Hunnius, 2015).

Clearly children must begin to exhibit motor resonance in response to acquired action words at some point within the learning process. To date, motor resonance induced by linguistic material in toddlerhood has not been investigated to the same extent as action observation. Therefore, the aim of the present study was to investigate the early link between language and motor systems. Motor activation can be assessed by means of EEG with the focus on mu rhythm, which reflects reduced power in the alpha frequency band over the sensorimotor cortex and which can be observed during action observation and execution (mu desynchronization). In adults, this frequency band typically ranges from 8 to 13 Hz compared to that of infants (6-7 Hz at 5 months), toddlers (8 Hz at 24 months), and pre-school children (9 Hz at 4 years of age) (Marshall, Bar-Haim, & Fox, 2002; Marshall & Meltzoff, 2011; Saby & Marshall, 2012). Importantly, however, studies comparing mu rhythm modulation to verbal and

non-verbal as well as visual material as a function of training, have to our knowledge, not been conducted.

In the present study, we were specifically interested in examining the modulation of the mu frequency band before and after observational experience of new actions that were consistently coupled with a novel verb or a novel non-action sound. We proposed that this novel pre-post within-subject training paradigm would give new insights into the relation between motor activation and language acquisition. Our study is therefore composed of a pre-training EEG assessment, an observational training phase, and a post-training EEG assessment. The primary focus of this study was to investigate if a) *verb comprehension would result in motor activation as indicated by modulation of mu rhythms over frontal and central sites as a function of observational learning experience.* Equally, we aimed to investigate whether b) *the suppression of mu in response to action-related auditory verbal material can be distinguished from that of auditory non-verbal material* (similar to the rattle in Paulus et al.'s, 2012 study). Modulated motor activation as a result of verb processing in this study would indicate a close functional involvement of both action and language systems in this task. In fact, the pattern of activation c) *may parallel the response to a familiar action verb at post-test assessment. Alternatively, however, we may find differences in activation between familiar and newly learned verbs.* Unlike the familiar control action, the presented novel actions were never actively experienced. Mechanisms between the two cerebral hemispheres in early development are currently unknown. It is possible that the processing of verbs currently outside ones' motor repertoire (similar to more abstract verbs in Rüschemeyer et al's., 2007 study) would be indeed more right lateralized when compared to verbs within the toddlers' motor repertoire. Finally, if activation of motor areas is indeed specific to mechanisms of action and language processing, d) *non-verbal*

referents presented alone should not lead to similar effects at post-test assessment. Similarly, we predicted e) *the suppression of mu to action-related visual and audio-visual material to be distinguishable between and after training as well as a function of referent (verbal vs. non-verbal referent).*

4.2 Methods

4.2.1 Participants

The final sample consisted of 20 monolingual English-speaking 2-year-olds (11 male) with an average age of 24 months (*SD* = 1.63) and 18 days (*SD* = 8.93) on visit 1 and 25 months (*SD* = 1.63) and 13 days (*SD* = 9.05) on visit 2. All toddlers were born full term (37-42 weeks gestation), were of normal birth weight (>2500g) and were recruited from the Birkbeck Babylab (CBCD) database in and around London, UK. Each participant's legal guardian received a detailed description of the study verbally and in the form of a participant information sheet prior to participation. Formal written consent was attained prior to participating. Each family received travel reimbursements for each attended visit (maximum of two visits) and a small gift in the form of a Birkbeck Babylab t-shirt or a shopping bag. The design of this paradigm required usable data from both time points in order to assess training effects in a within subject manner. As a consequence, participants (*n* = 3) were tested but not invited back due to their lack of attention to the stimuli at pre-training assessment and participants (*n* = 4) had to be excluded from the final analysis due to their inattentiveness on the subsequent visit (post-training assessment). Finally, data from four further participants were excluded from the final dataset due to technical problems (*n* = 3) or watching the wrong training material at home (*n* = 1).

4.2.2 Stimuli

The experimental stimuli comprised auditory, visual and audio-visual action stimuli. At pre-training assessment (visit 1) toddlers saw one familiar and two novel stimuli. At post-training assessment (visit 2) the stimulus material consisted of one familiar, two trained and two additional novel auditory stimuli. In addition, two novel actions and one familiar action were featured in the visual, & audio-visual modality at pre-test and the two novel actions were also presented at post-training assessment. A block design was chosen when presenting the different modalities with a hierarchical order starting with *Block 1* for the most abstract auditory action stimulus material, *Block 2* representing the visual action material, *Block 3* displaying audio-visual stimuli and *Block 4* depicting familiar action material in a mixed modality fashion. The blocks differed between the pre- and post-training assessment in some respect due to the inclusion of the two additional novel stimuli (*Block 1*) and a complete exchange of *Block 4*. A detailed description is given below.

4.2.2a Auditory Stimuli

One novel action verb *yokking* served as the linguistic auditory stimulus. To evaluate whether verb processing specifically induces motor activation, rather than a more general mapping of a visual event to a referent a second auditory stimulus, a two tone piano sequence, here referred to as *Sound 1,* served as a non-linguistic auditory stimulus. Both novel auditory stimuli were of 800 ms duration with an inter-stimulus interval (ISI) ranging from 1200-1600 ms selected pseudo-randomly such that neither the verbal nor the non-verbal referent was presented more than twice consecutively. To facilitate language encoding, a native English female speaker from the South of England provided the linguistic cue, congruent with the accents typically encountered by the participants. To keep toddlers' attention focused on the screen during the

auditory presentation in both pre- and post-training assessment, a still image (a hidden object scene, with content unrelated to our paradigm) was presented in the centre of the screen for the full duration of the auditory presentation. Within *Block 1* at pre-training assessment, the two auditory stimuli could be presented 8 times each (see Figure 4.1).

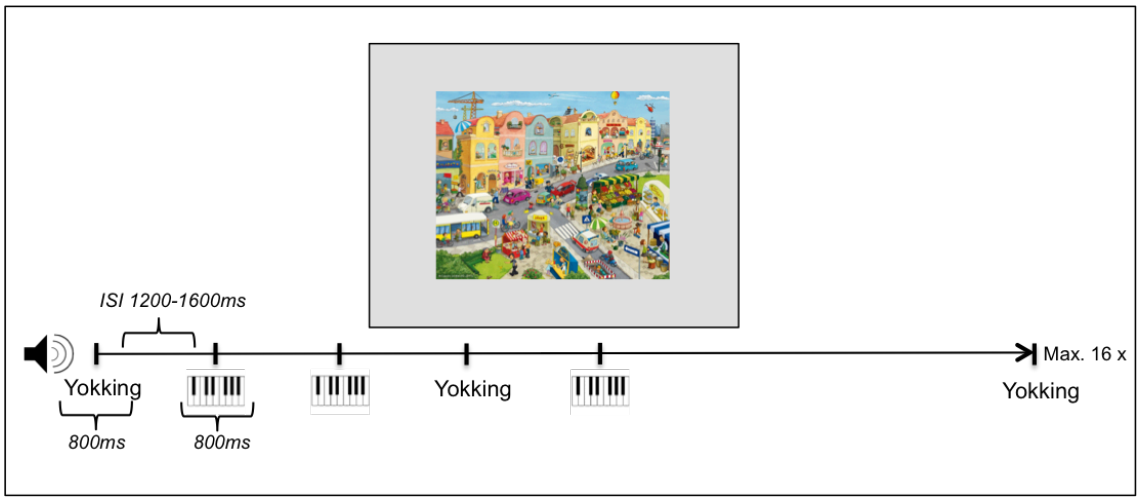

Figure 4.1: *Auditory presentation (Block 1) of the pre-training assessment. A still image was presented in the centre of the screen for the full duration of the auditory presentation of the novel verb and sound.*

In addition to the novel stimuli, one familiar action verb *squeezing* was chosen, with the aim to assess whether mu rhythm of 2-year-olds is indeed modulated in response to familiar action verbs, as in adults (Moreno, Vega, & Leon, 2014). This auditory stimulus was again 800 ms in duration with an ISI ranging from 1200-1600 ms and presented in a separate block (*Block 4*) with a total of 18 presentations in a mixed modality fashion, meaning that during a complete block, a child could be presented with 6 auditory, 6 visual and 6 audio-visual stimuli concerning the squeezing action with the constraints that one particular modality was not presented more than two times in a row.

During the second visit (post-training assessment), two additional auditory stimuli were presented alongside the now trained auditory stimuli from visit 1. The two additional novel stimuli comprised a novel verb *tamming* again spoken by the same

female voice and a second novel sound consisting of a different set of two piano tones (Sound 2). Stimulus duration was matched to all previously presented auditory stimuli (800 ms). Block 1 at post-training therefore differed from Block 1 at pre-training in the number of auditory stimuli presented. Instead of just the two stimuli (yokking and Sound 1) we now also included the two novel stimuli (tamming and Sound 2) as well as the familiar word (squeezing). All 5 auditory stimuli were presented for a maximum of 25 presentations (5 x yokking, 5 x tamming, 5 x squeezing, 5 x Sound 1, 5 x Sound 2) with the constraints that verbal or non-verbal referents were not presented more than 2 times consecutively. The same picture was presented to the participants as in Block 1 (at visit 1) in order to retain the toddler's attention during the auditory presentation with the ISI between auditory presentations of 1200-1600 ms, jittering at random (see Figure 4.2).

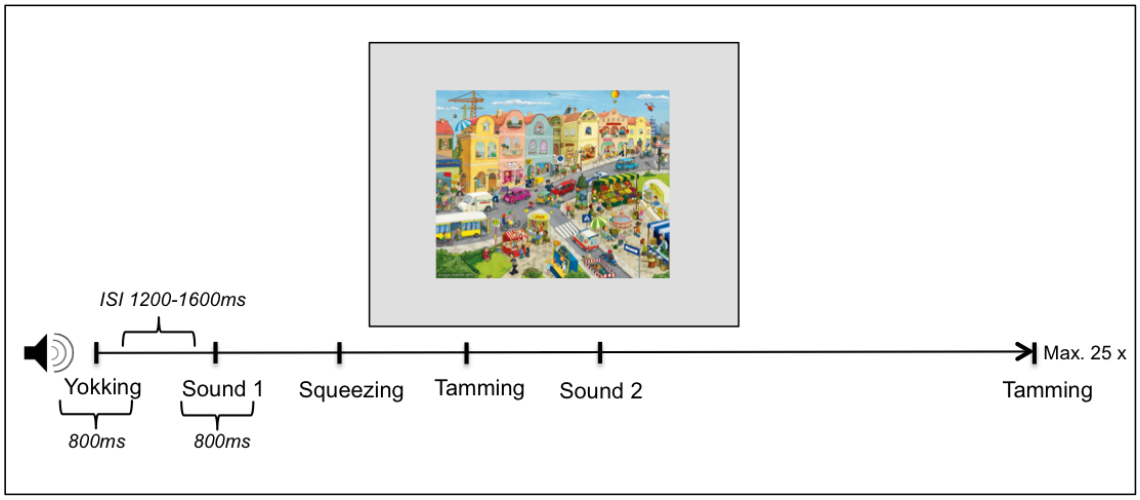

Figure 4.2: *Block 1 at post-training assessment. Participants saw the same still image in the centre of the stimulus monitor during the presentation of the two trained, two novel and 1 familiar auditory stimuli. Within one completed block, the five auditory stimuli could be presented 25 times in total.*

4.2.2b Visual Stimuli

Manual actions have been previously identified as significantly modulating action-related frequency bands in adults (e.g., Moreno et al., 2013). Therefore, actions

performed in the present stimulus material are executed with one hand. The experimental visual stimuli comprised three short action video clips (1 familiar action squeezing and 2 novel actions) each of which was 3900 ms in duration. All actions were performed with the same medium kinetic sand, which has unique textural properties, resulting in a novel effect of prolonged movement of the sand even after finishing the manipulation upon the sand. Each action was presented equally often from the left and right side. The video clips always started with presenting the hand hovering over the sand for the first 600 ms. Initially the sand was presented moulded into an upturned cup shape, identical across stimuli. Thus, from the starting position it was impossible to anticipate which action was to follow. Subsequent to the initial hand in starting position (hovering over the sand), the action was initiated. Novel Action A was performed with the index finger starting from the top centre of the moulded sand, slowly pushing the finger down through and out from the sand making the sand split in the middle and causing the sand to move in the same trajectory. After action execution, the hand disappeared from the scene (either from the left or right depending on the orientation of the hand). Novel Action B differed from Action A in its usage of two fingers, namely the index finger and the thumb. The performed action started by first rotating the flat hand that was hovering over the sand at 90 degrees and forming a pincer grip with which the end corners of the moulded sand were picked up and dropped down. The sand then performed a self-propelled motion in the same trajectory and simultaneously the hand disappeared from the screen in a similar fashion as in action A. The familiar action was performed on the same medium and displayed squeezing the sand, this time with the use of the full hand. It is the only action that does not cause the sand to move, but similar to actions A and B, leaves the sand in a different end state following squeezing.

All actions were held equal with respect to the length of active manipulation of the sand (hand-sand contact), which was 2000 ms long for all action (see Figure 4.3).

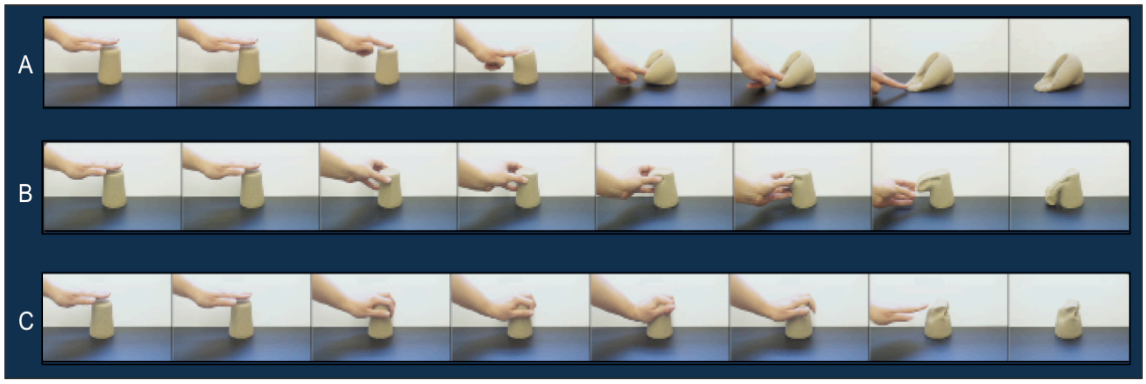

Figure 4.3: *The top two rows represent the two novel actions, action A - performed with the index finger, Action B - performed with the index finger and the thumb and action C - the familiar action squeezing. All actions involved manipulating the medium kinetic sand, presented from the left and right equally often. Timings for each of the video clips were the same. Each action started with a 600 ms period in which the hand hovered above the moulded sand before the action itself unfolded.*

Participants were presented with the visual stimuli in block format with the two silent novel actions (Action A & B) presented in *Block 2*. The identical two novel actions *but* presented with an auditory referent were presented in *Block 3*. The familiar action was presented in the final *Block 4*, which was a mixed modality fashion (auditory only, visual and audio-visual). *Block 2 & 3* remained the same for the post-training assessment whereas the familiar action was no longer present.

4.2.2c Audio-Visual Stimuli

The same action video clips as described in section 4.2.2.b (Familiar and Novel actions A & B) were now accompanied by the auditory referents (verbal referents - yokking & squeezing as well as Sound 1). The video clips could therefore be further subdivided into three distinct time windows. Six hundred milliseconds into the video for instance, either the verbal or non-verbal referent was presented together with the start of the hand

action leading to one of the three actions. This time window was 800 ms of duration, which represents the length of the auditory stimulus (TW1 600-1400 ms). The second time window is characterized by the offset of the auditory stimulus and the hand was now in contact with the precast sand (TW2 1400-2000 ms). The subsequent and final time window (TW3 2000-2800 ms) represented the action execution stage and was unaccompanied by any additional auditory stimulation. As described previously, the novel audio-visual stimuli were a feature of *Block 3* at both time points (pre- and post-training). The familiar audio-visual action was only presented in the initial pre-training assessment in the mixed modality *Block 4*. The novel verb and sound action pairs were counterbalanced across the sample so that for half of the group, novel action A constituted the critical action word *yokking* and the novel *Sound 1* referred to the novel action B. For the remaining children, novel action B was most critical for the novel word *yokking* and action A for the novel action *Sound 1*. Therefore, any unintentional differences in salience between the two visual actions were intentionally minimised.

4.2.2d Training Videos

The training video was 2.38 minutes long (including a short 4 s greeting at the start of the training and a 4 s daily comment reminder at the end of the training session) and comprised of 32 actions (16 x Action A, 16 x Action B) in a pseudo-random order with the restriction that the same action was not presented more than 2 times consecutively. The actions were performed from the left and right equally often and the actions itself varied in length (M = 4.52 s, range: 3.1 s - 5.7 s). Sound-action pairings were counterbalanced across participants. The sequence of the training video was the same between the two training groups. To further maintain toddlers' interest and to facilitate word learning by means of multiple exemplars (e.g., Twomey, Ranson, & Horst, 2014), both novel action A & B were performed with different people's hands, at different

angles, in front of various backgrounds. The verbal referent was always introduced with an introductory sentence *Look! I'm yokking* (1600 ms duration) spoken by the same female voice as for all verbal auditory stimuli presented in the experiment. The second novel action was similarly introduced with a longer piano tone sequence (1600 ms) with the two critical tones presented at the end of the sequence similar to the verb *yokking* being presented at the end of the sentence.

4.2.3 Procedure

4.2.3a Pre-Training Assessment

After the warm up phase in the testing cabin, children were seated on their caregivers lap in a dimly lit, soundproof and electrically shielded experimental cabin in front of a 20-inch stimulus monitor. The room was further outfitted with a camera, microphone and loudspeakers. EEG was recorded continuously with a 128-channel Geodesic Sensor Net (HCGSN 130; EGI, Eugene, OR). After the EEG net had been correctly fitted, the experimenter left and closed the cabin, started the experiment and monitored the testing session from outside the testing area via the experimental monitors. During the experiment, each child was first presented with *Block 1*, followed by *Block 2, 3* and *4*. In the event that the child increased their body and head movement due to the gradually decreasing attention to the experimental stimuli, the experimenter had the freedom to terminate the currently presented block by a button press, which resulted in the immediate presentation of the next consecutive block. In this way, the child's attention could be brought back to the experimental monitor, reducing body and head movements. This procedure could be repeated with any given block. The experimenter also offered breaks in appropriate moments. For the duration of the pause the monitor remained dark. The experiment ended once the toddlers' attention could no longer be maintained. The participant's behaviour was video-recorded throughout the session for

offline coding of looking behaviour to the experiment and to control for the amount of motor activity and vocalization in each condition. After the experiment, those children who attended to the stimuli and generated sufficient data in the first data acquisition phase were invited back for the second stage of the study. Caregivers were informed prior to consenting to take part in the study that there was a possibility that there may not be a second visit to the laboratory as the toddlers' first EEG session served as an individual baseline (pre-training measure) - a crucial part of this study design. The aim was to ensure that caregivers and their children did not have to complete training and return to the laboratory for a second EEG assessment unless necessary. For those who were invited back, the experimenter explained the home training procedure in detail and caregivers were given a home training instruction sheet to take home. Children were further assessed for their ability to perform the squeezing action on the first day of their visit by asking the caregiver if their child is familiar with this action as well as asking the child directly to squeeze a teddy bear's paw in a child directed way.

4.2.3b Home-Training

The training typically started one day after the first visit to the laboratory and lasted 6 days on average (range: 5-7). However, for some parents it was not possible to return within a week's time. When longer breaks between visits were unavoidable, the experimenter reminded families of the training start 7 days prior their return to the laboratory. Parents could choose the time of day in which they performed the training but were asked to take into consideration that the training should be performed during a time in which their child was most likely to be attentive and most awake. Further, we asked parents to reduce external distractions and to make notes of the quality of the training session daily. The training videos were viewed on a computer, laptop or tablet device but we ensured that smart phones were not used due to the small screen size. To

monitor training commitment and the child's attention to the training videos, each parent was given an unique participant ID number and log in details for an email account, which was specifically set up for the purpose of this study. Parents were asked to sign in to this account every day to retrieve the training video YouTube link. The email that contained the link further contained instructions on how to complete the training and leave daily comments. The parents had to click on the link provided in the email while remained signed in. This way when leaving comments, no personal details were displayed with the comments. The training video had to be viewed for the full length of the training session and immediately after the training parents were asked to state on a Likert scale from 0 to 5 (with 5 meaning full attention to the training video) how their child attended to the training video for each day. Toddler's mean attention across the duration of the training was rated with 3.4 (range: 1.7 - 4.8). Further, parents were asked to indicate the date, time and additional notes if applicable. Such additional notes could be anything that parents have observed before, during or after the training such as the child repeating the word or any attempts to imitate the actions. This procedure further helped to monitor that the training had been performed correctly. Parents were invited back for the post-training EEG assessment on the 7th day – or one day after the last training session. Lastly, parents were requested not to practice the actions or mention the word within or outside the training time.

4.2.3c Post-Training Assessment

The EEG procedure in this final phase was similar to that of the pre-training assessment with the exception of Block *1 & 4*. In *Block 1*, as described above, 3 auditory stimuli were presented (namely yokking, squeezing and Sound 1) together with two additional novel auditory stimuli as controls. These consisted of an additional novel verb tamming and a non-verbal referent Sound 2. The familiar mixed-modality Block 4 of the pre-

training assessment was further replaced by an ERP Block with the aim to examine toddlers' semantic knowledge by means of measuring any produced N400 component to matches and mismatches of the auditory referents for the critical mid state of the two now trained actions. Van Elk et al. (2010) observed a significant correlation between the N400 ERP component and mu power in the sense that motor activation to action language was inversely related to the size of the N400 component in adults, supporting the view that motor resonance plays a functional role in semantic integration. It should be noted that this additional ERP block did not result in a large enough sample size for analysis, and is therefore only briefly mentioned in this experimental chapter in order that the wider experimental paradigm and context of the study is conveyed. All four blocks of the post-training assessment always followed the Block 1-4 hierarchical orders and the experiment would loop to Block 1 after completion of the last block until the toddler's attention could no longer be maintained (see Figures 4.4).

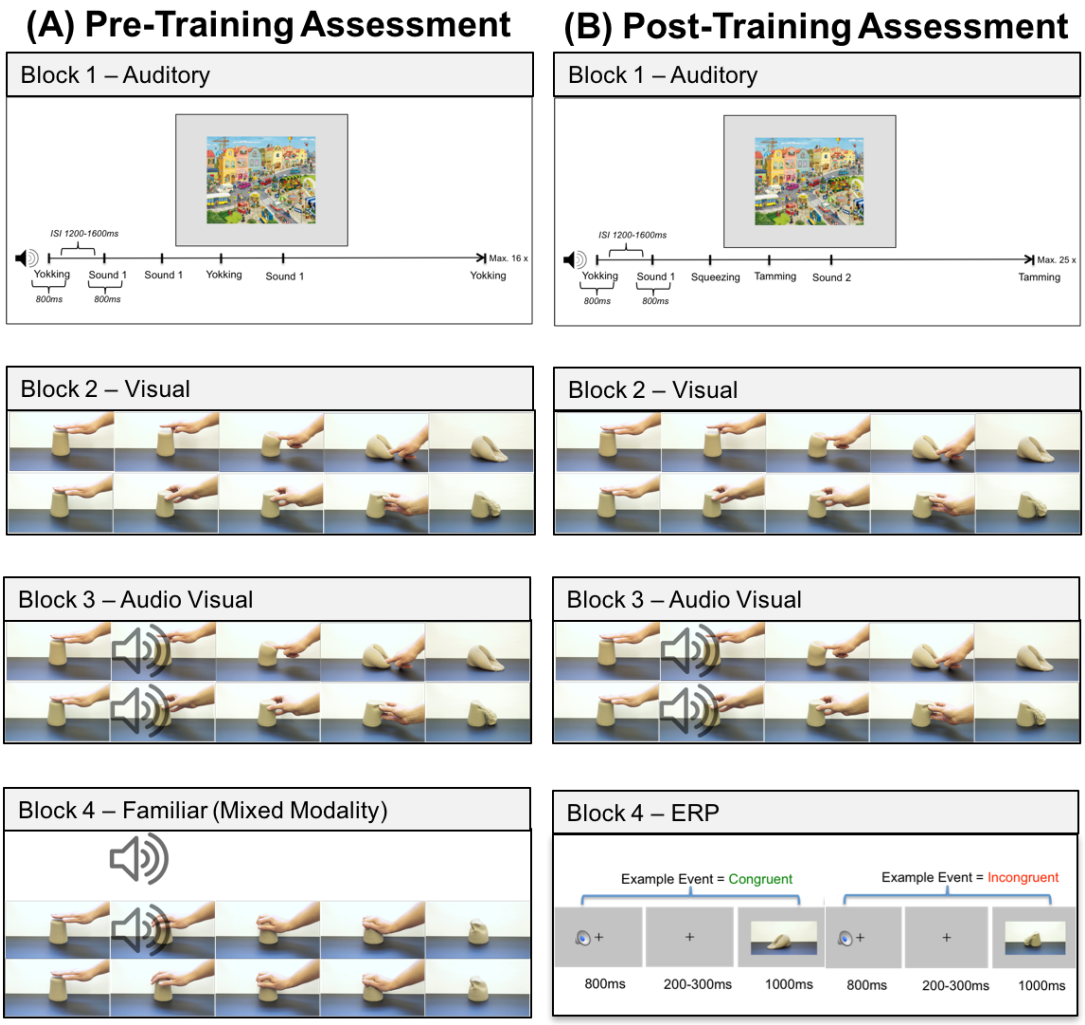

Figure 4.4: *Left depicts the four blocks presented at pre-training and right the four blocks presented at post-training EEG assessment.*

4.2.4 EEG recording and analysis

EEG was recorded via 128-channel Geodesic Sensor and amplified with an EGI Net Amps 400 amplifier with a sampling rate of 500 Hz (HCGSN 130, EGI, Eugene, OR, USA). EEG was referenced online to the vertex electrode (Cz), and an analogue bandpass filter was applied (0.1 to 100 Hz) before data was stored for subsequent off-line analysis. Data for auditory stimuli were segmented with a 1000 ms baseline and a 1500 ms period post stimulus presentation. For the visual stimuli the segments comprised of a 900 ms baseline period followed by a 4500 ms post presentation phase.

Subsequently, video-recordings were coded for toddlers' attention to the stimulus material for post processing. Remaining attended trials were manually and visually inspected for artifacts caused by eye or body movements with segments including more than 10 channels with excessive noise levels being rejected. Noisy channels from the remaining data, which comprised fewer than 10 bad channels, were interpolated by using the average of surrounding electrodes. EEG data was post processed using custom-made scripts in *WTools* (available on request) and EEGLab (9.0.5.6b). The time-frequency analyses were conducted performing a continuous wavelet transformation. Complex Morlet wavelets were computed at 1 Hz frequency intervals for the frequency range 4-15 Hz. Total spectral activity was calculated performing convolutions with the wavelets on all channels. The absolute value of the results was computed and served as the dependent variable. The transformed epochs were averaged for each condition (see Csibra et al., 2000). Finally, 400 ms at the beginning and the end of each segment for the auditory stimuli and 700 ms for the visual stimuli were removed to avoid distortion due to the transformation. Baseline correction was performed at each frequency by subtracting the mean activity of 200 ms before stimulus onset from the signal. The grand average was calculated for both conditions and time points separately.

4.2.4a Analysis considerations

For the auditory domain, 20 participants contributed sufficient trials for the initially novel (pre-training), subsequently trained verbal and non-verbal auditory referents (post-training) as well as the control stimulus, squeezing. If we considered including those participants that contributed enough artifact-free trials for all 5 auditory stimuli at post-training, our sample size would have reduced to 11 participants due to the demands of the study. Consequently, and as our primary focus was to investigate motor modulation before and after training for yokking and Sound 1, we rejected analysis for

the two additional novel auditory referents at post-training. Further, *Block 2* (visual only) from pre- and post training assessment as well as *Block 4* (ERP) of the post-training assessment failed to establish a large enough sample size (n = 10) to pursue subsequent analysis and were also excluded from further consideration. The subsequent table describes the mean number of trials and standard deviations for the included independent variables of this study as well as the mean and standard deviations of total trials presented (see Table 4.1).

Table 4.1.

Mean and Standard Deviation (including the total number of presented trials) and sample size for the auditory and audio-visual task separate for the pre- and post-training assessments.

	Auditory								
	Pre-Training					*Post-Training*			
	N	Mean (Total)		SD (Total)		Mean (Total)		SD (Total)	
Sound 1	20	9.7	20.5	3.9	3.6	20.5	40.7	10.0	13.8
Yokking	20	9.9	20.4	4.0	3.7	9.9	16.7	3.9	5.0
Squeezing	20	4.9	10.5	3.9	3.6	5.1	8.3	2.1	2.7
Sound 2	11	-	-	-	-	5.8	9.3	1.8	2.4
Tamming	11	-	-	-	-	6.1	9.2	1.4	2.8
	Audio-Visual								
	Pre-Training					*Post-Training*			
	N	Mean (Total)		SD (Total)		Mean (Total)		SD (Total)	
Sound 1	14	8.7	19.9	4.4	5.6	6.2	13.1	3.9	6.1
Yokking	14	8.9	19.6	3.9	5.1	6.2	13.2	4.1	4.7
Squeezing	14	5.6	10.8	1.9	2.7	-	-	-	-

4.3 Results

4.3.1 Motor Modulation in response to Familiar, Novel & Trained Auditory Material

We hypothesised that verb comprehension would result in motor activation as indicated by modulation of mu rhythms over frontal and central sites as a function of

observational training. The suppression of mu in response to action-related auditory verbal material should further be distinguished from that of auditory non-verbal material. Motor activation in response to trained verbs at post-test assessment is expected to parallel the response to a familiar action verb at pre-training assessment but differences in activation between familiar and newly learned verbs can still occur due to the lack of active training experience.

Present literature on sensory motor alpha by Marshall and colleagues (2002) report a dominant peak frequency for the majority of toddlers aged 24 months at 8 Hz. This was backed by visual inspection of the current dataset, which led us to consider this particular frequency for subsequent analysis of the auditory data set. Hemispheric differences are further reported for sound and language processing in adults (e.g., Rüschemeyer et al., 2007), which was also evident in the present dataset. Based on these assumptions and additional visual inspection of the 3D topographical maps of the current auditory dataset, hemispheric activation was observed between the familiar (*squeezing*) and trained verb (*yokking* - post-training) with the latter being more activated over right central sites in comparison to the left central activation for the familiar action verb in the observed time window. The latency and frequency band was more pronounced for the familiar verb starting at the onset of presentation and displaying a wider frequency band of 6-9 Hz when compared to the trained verb at 8 Hz. In order to directly compare the experimental conditions, we chose the time window of 600 - 800 ms post stimulus presentation at 8 Hz for the following 5 left central (30, 31, 36-C3, 37, 54) and 5 right-central (79, 80, 87, 104-C4, 105) electrodes (Figures 4.5 & 4.6).

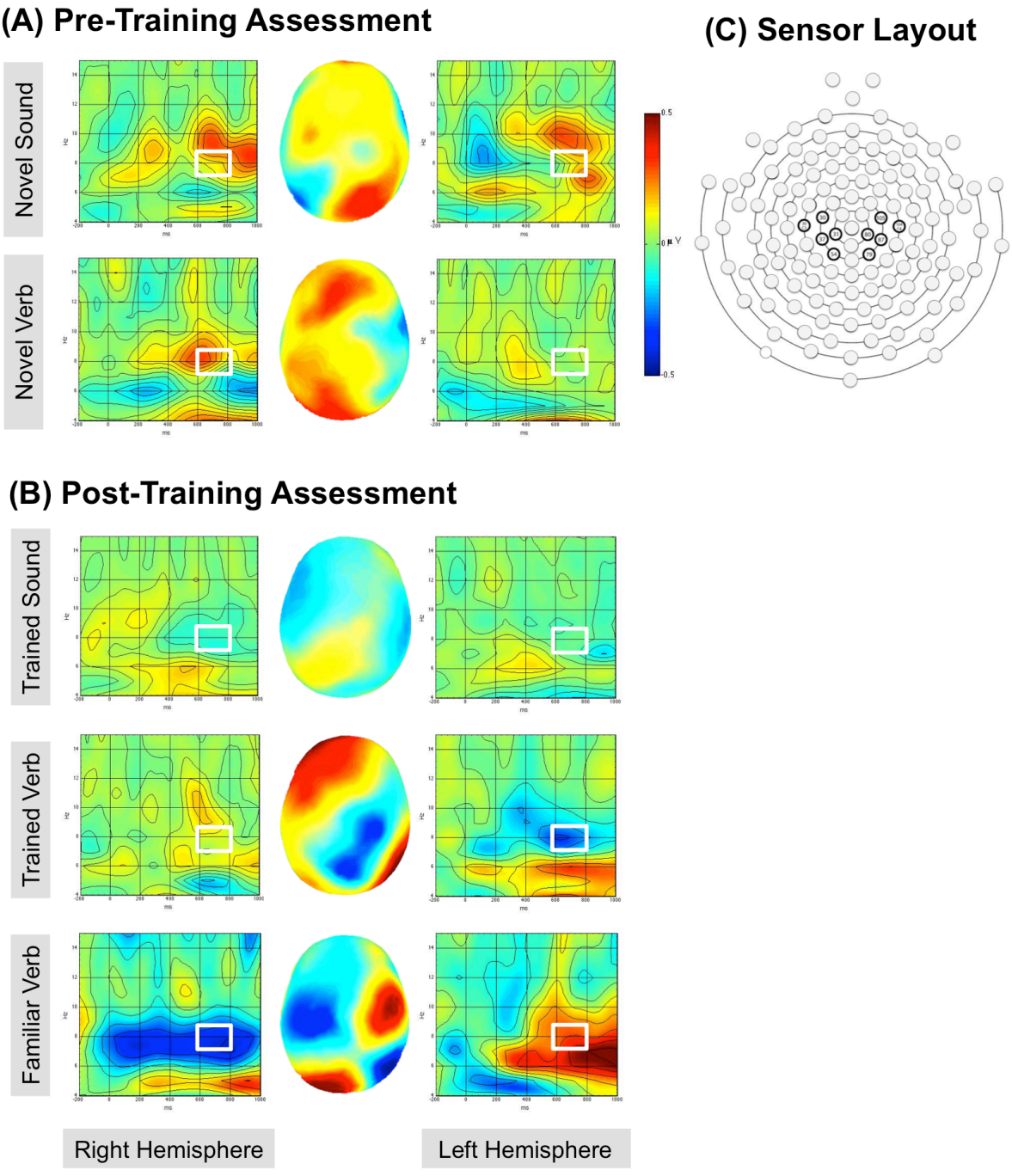

Figure 4.5: *The top two rows depict activation for the perceived auditory sound and novel verb at pre-training assessment (A) and the three subsequent rows depict activation at post-training (B) for the now trained sound, trained verb and the control variable familiar action verb for the entire segment (-200 – 1000 ms) separate for the left and right hemisphere. Corresponding 3D Topographical Scalp Maps averaged for the analysed time window (600-800 ms) across all participants (N= 20) at 8 Hz are displayed in the middle. Area of channels included for the left and right are marked on the sensor layout to the left (C).*

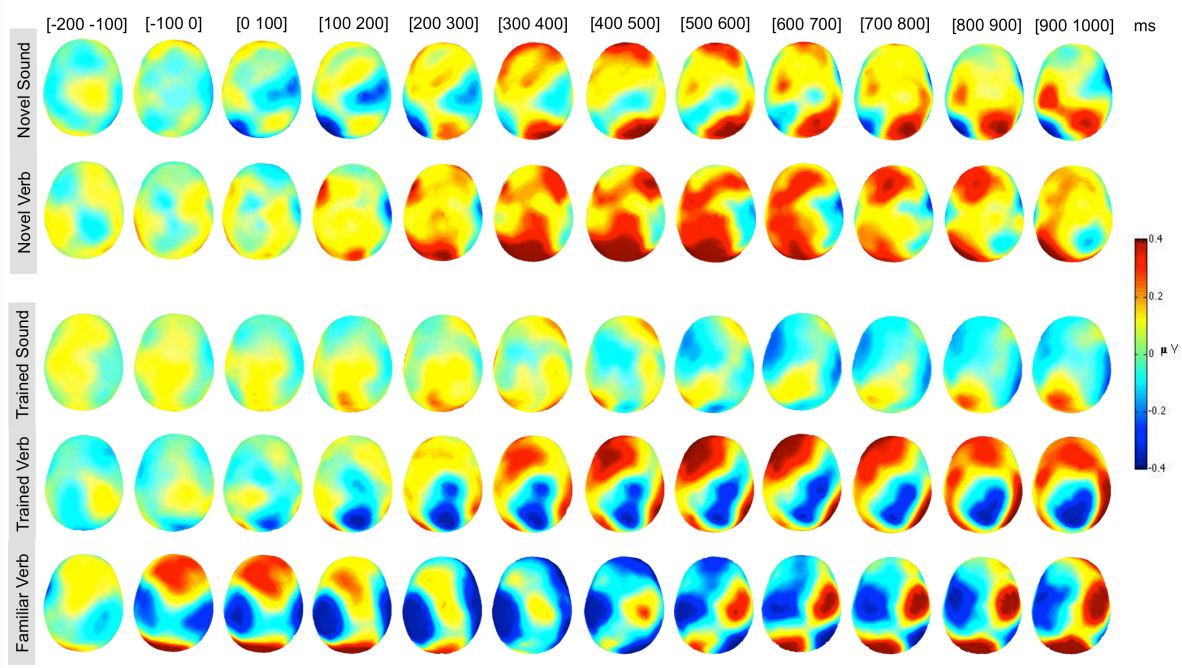

Figure 4.6: *3D Topographical Scalp Maps across all participants (N= 20) in 100 ms intervals for the segment length including a 200 ms baseline and post stimulus activation for 1000 ms duration post stimulus presentation at 8 Hz. The top two rows depict activation for the perceived auditory sound and novel verb at pre-training assessment and the three subsequent rows depict activation at post-training for the now trained sound, trained verb and the control variable familiar action verb.*

To test our primary hypothesis if novel auditory material will result in motor activation as a function of training and to test for any distinguishable differences in motor desynchronization between verbal and non-verbal referents, we first performed a 2 (Condition - Sound1/Novel Verb) x 2 (Time Point - Pre/Post-Training) x 2 (Hemisphere - Left/Right) repeated measures ANOVA. The analysis revealed a significant Condition x Hemisphere interaction, $F(1, 19) = 4.657$, $p = .044$, partial $\eta^2 = .197$. No other main effects or interactions yielded significance.

Although there was no 3-way interaction between Condition, Time Points, and Hemispheres, we expected there to be differences in activation between the sound and language conditions at post-training assessment. Therefore, we next performed a 2 (Condition - Sound1/Novel Verb) x 2 (Hemisphere Left/Right) repeated measures

ANOVA separately for each of the two time points (Pre/Post-Training). For the pre-training stage, we found no significant main effects or interaction, indicating that the interaction must be driven by the post-training assessment. Results from the post-training analysis revealed a significant Condition x Hemisphere interaction, $F(1, 19) = 6.357$, $p = .021$, partial $\eta^2 = .251$. Follow-up paired sample t-tests (two-tailed) revealed significant differences in alpha power for the Trained Verb between the left and the right hemisphere, $t(19) = 2.836$, $p = .011$, $d = .669$ ($M_{left} = .084$, $SD = .630$, $M_{right} = -.289$, $SD = .474$), as well as a significant difference between the Trained Verb and Trained Sound over right central channels, $t(19) = -2.321$, $p = .032$, $d = .626$ ($M_{trainied\ verb} = -.289$, $SD = .474$, $M_{trained\ sound} = -.026$, $SD = .358$).

Next, to explore differences in motor activation within the post-training assessment for the trained Sound and Verb with an already familiar action-referent, we performed an additional 3 (Condition - Trained Sound/Trained Verb/ Familiar Verb) x 2 (Hemisphere - Left/Right) repeated measures ANOVA which yielded a significant Condition x Hemisphere interaction, $F(2, 38) = 7.574$, $p = .002$, partial $\eta^2 = .285$ with the two main effects not reaching significant levels. Follow-up paired sample t-tests revealed significant differences between the left and right hemisphere for the Familiar Verb, $t(19) = -2.644$, $p = .016$, $d = 1.030$ ($M_{left} = -.381$, $SD = .669$, $M_{right} = .293$, $SD = .639$). For the right hemisphere, the pairwise comparison further revealed significant differences between the Trained Verb and the Familiar Verb, $t(19) = -2.680$, $p = .015$, $d = 1.034$ ($M_{trained\ \text{verb}} = -.289$, $SD = .474$, $M_{familiar\ verb} = .293$, $SD = .639$). Trained and Familiar Verbs were also significantly different over left channels, $t(19) = 2.210$, $p = .040$, $d = 1.871$, with the familiar verb eliciting more negativity ($M_{familiar\ verb} = -.381$, $SD = .669$) compared to the trained verb ($M_{trained\ verb} = .835$, $SD = .630$).

Finally, to establish whether motor activation differed significantly from baseline, one-sample *t*-tests (with the test value = 0) were performed for all conditions and time points, separately for the two hemispheres. The results indicated that only the conditions Trained Verb on the right hemisphere, $t(19) = -2.726$, $p = .013$, and Familiar Verb on the left hemisphere, $t(19) = -2.545$, $p = .020$, differed significantly from baseline (see Figure 4.7 for a summary of all significant results and the power relative to baseline)

Figure 4.7: *Mu power for the left and right hemisphere separately over central electrodes to the familiar verb squeezing (grey), novel sound (blue), and novel verb yokking (red) pre-training (solid colour) and post-training (textured). Note the familiar verb derived from Block 1 at post-training. (** $p < 0.01$, * $p < 0.05$ - green highlights significance from baseline, black indicates significance between conditions).*

4.3.2 Motor Modulation in response to Familiar, Novel & Trained Audio-Visual Material

Our hypothesis in the present study is that mu suppression to action-related visual and audio-visual material can be distinguished before and after training and as a function of referent (verbal vs. non-verbal referent).

The audio-visual material in the present dataset recruited a larger fronto-central area leading us to choose an array of channels representative of the primary motor region (5, 6, 12, 13, 20, 28, 29, 34, 35, 39, 40, 44, 109, 110, 111, 112, 114, 115, 116, 117, 118). Due to dissimilarities between the three time windows (e.g. TW 1 being the only time window representing auditory and visual cues), we expected variation in activation across the time windows. Based on these criteria and further visual inspection of the 3D scalp maps for each of the three time windows and activation averaged across the 21 channels of interest, we identified mu suppression to be most representative at 6-8 Hz (Figure 4.8).

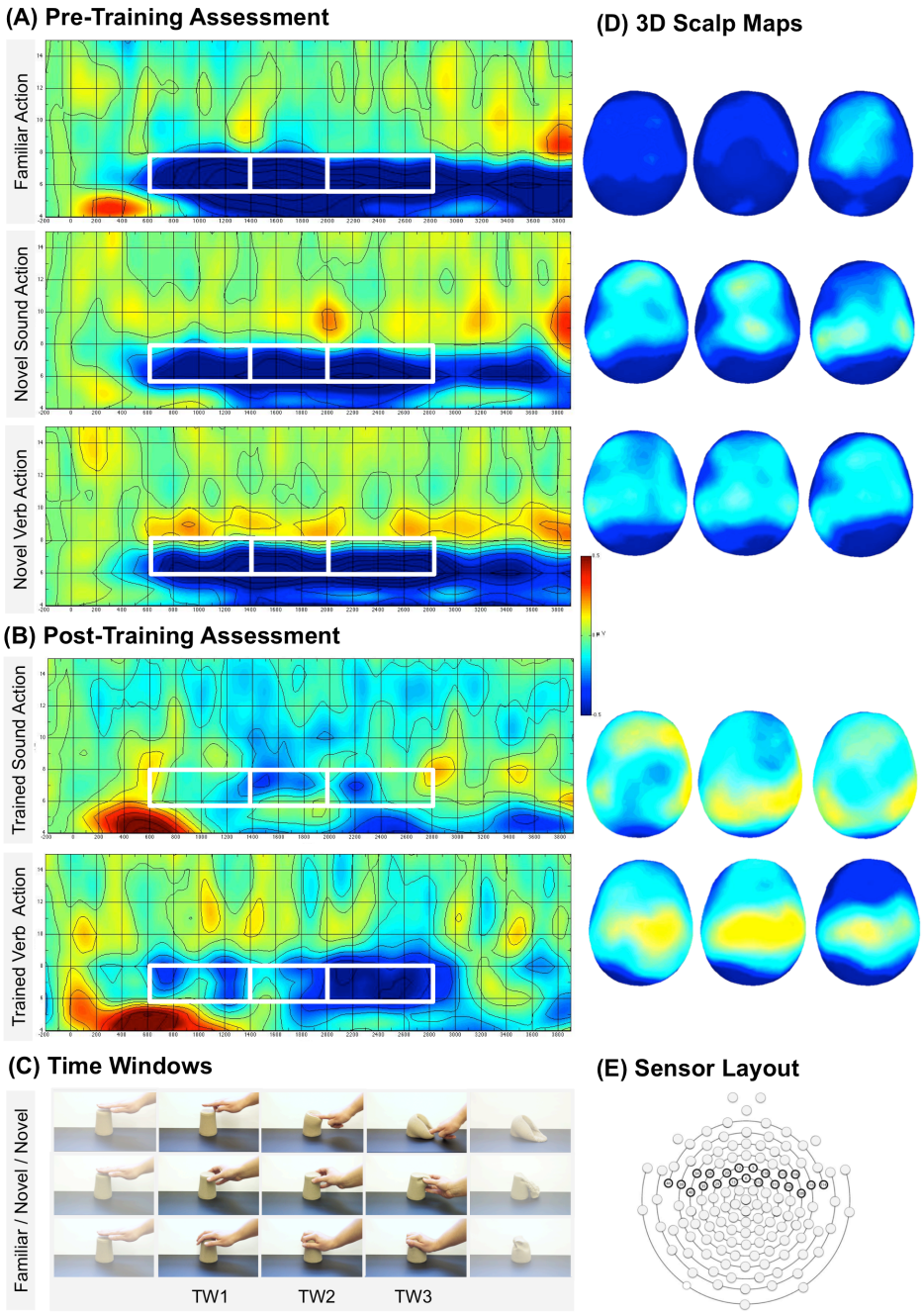

Figure 4.8: *The top three rows depict activation for familiar and two novel audio-visual materials and pre-training (A) and the two subsequent rows depict activation at post-training for the now trained language and non-language video material (B) for the entire segment (-200 – 3900 ms). Analogous of the three distinct time windows (C): TW1 (600-1400 ms - auditory cue), TW2 (1400 – 2000 ms - offset of the auditory stimulus and first hand to sand contact), and TW3 (2000-2800 ms - action execution). (D) Corresponding 3D Topographical Scalp Maps for each of the three time windows averaged for the analysed time across all participants (N= 14) at 6-8 Hz. Area of channels included for analysis are marked on the sensor layout to the bottom left (E).*

First, we examined our primary hypothesis of motor modulation to audio-visual verbal and non-verbal stimuli as a function of training. We performed a 2 (Condition - Sound1/Novel Verb) x 2 (Time Point - Pre/Post-Training) x 3 (Time Window 1/2/3) repeated measures ANOVA[9]. This analysis revealed no significant Condition x Time Window interaction, $F(2, 26) = 3.221$, $p = .056$, partial $\eta^2 = .199$. Further, within the three main factors, no significant levels were reached.

Stronger premotor activation for non-familiar actions when compared to more familiar ones (e.g., Voght et al., 2007) were also observed in the present data before training. To account for this phenomenon we nevertheless performed two separate repeated measures ANOVA's for each of the recorded time points (Pre- and Post-Training). The pre-training assessment analysis failed to yield significant differences similar to the auditory domain. At post-training assessment, the Condition x Time Window interaction reached significance, $F(2, 26) = 3.597$, $p = .042$, partial $\eta^2 = .217$. Follow up paired sample t-tests revealed that the significant interaction was driven by differences in alpha power within the third Time Window, $t(13) = 2.980$, $p = .011$, $d = .691$, with actions trained with the language referent inducing more alpha suppression ($M_{trained\ verb} = -.407$, $SD = .454$) when compared to actions trained with the non-language referent ($M_{trained\ sound} = -.094$, $SD = .452$).

Next, we explored possible differences in motor suppression between the familiar and novel audio-visual presentations. A 3 (Condition – Novel Sound/Novel Verb/ Familiar Verb Action) x 3 (Time Window – 1/2/3) repeated measures ANOVA. This statistical analysis revealed a significant main effect for condition $F(2, 26) = 4.002$, $p = .031$, partial $\eta^2 = .235$ as well as a significant main effect for Time Window, $F(2,$

[9] We also considered hemisphere as a factor to enter the analysis, however the visual inspection of the dataset did not suggest hemispheric differences, and we did not find significant differences between hemispheres.

26) = 4.145, p = .027, partial η^2 = .242. The post hoc t-test (Bonferroni corrected) revealed that the significant main effect for Condition was driven by significant differences between the Novel Verb Action and the Familiar Verb Action, (p = .006), ($M_{novel\ verb\ action}$ = -.387, SD = .695, $M_{familiar\ verb\ action}$= -.650, SD = .673). Further comparisons (Bonferroni corrected) revealed the second Time Window was significantly different from the third time window, (p = .038) (M_{TW2}= -.552, SD = .705, M_{TW3}= -.461, SD = .635).

Since Time Window 1 was the only one in which both auditory and visual information were presented, this time window was of particular interest in order to fully explore the effect of the novel auditory referent in comparison to the familiar action verb at pre-training assessment. Thus, we performed a one-way repeated measures ANOVA with the three levels Novel Sound, Novel Verb and Familiar Verb. This analysis reached significance, F(1.55, 20.13) = 4.257, p = .037 (Greenhouse-Geisser corrected), partial η^2 = .247. The post hoc test (Bonferroni corrected) revealed that the significant main effect for Condition was driven by significant differences between the Novel Verb Action and the Familiar Verb Action, (p = .004), ($M_{novel\ verb\ action}$ = -.318, SD = .638, $M_{familiar\ verb\ action}$= -.628, SD = .644).

Lastly, we tested for motor activation against baseline. One-sample t-tests (with the test value = 0) were performed for all conditions and time points separately for the three time windows. For post training, only the actions trained with the language referent in the third time window significantly differed from baseline, t(13) = -3.353, p = .005. At pre-test assessment, all of the conditions differed significantly from baseline in all of the three time windows apart from the language referent condition, which did not for time window 1 and 2 (see Figure 4.9 for all significant differences between conditions and from baseline).

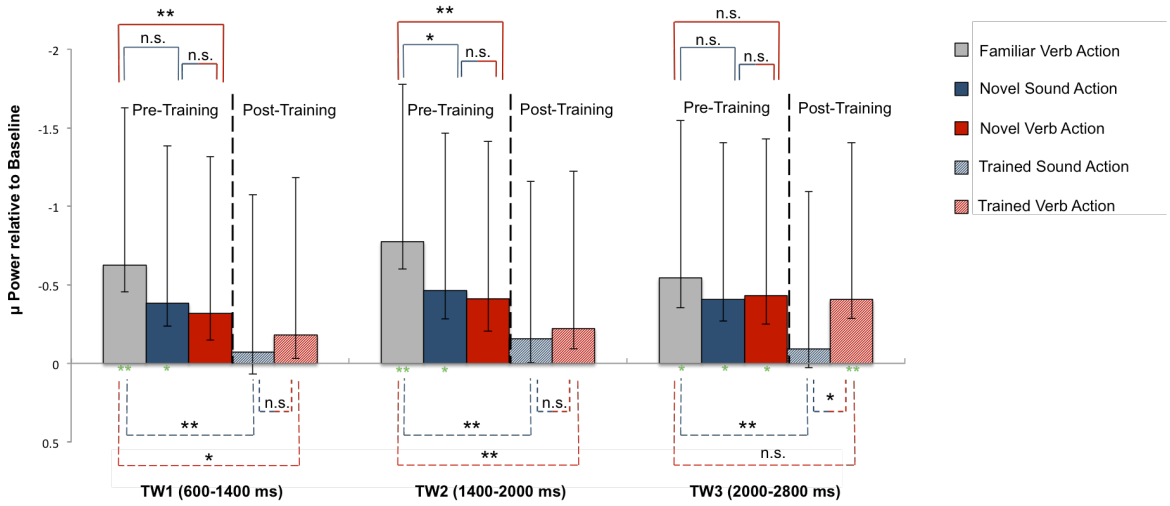

Figure 4.9: *Mu power to familiar, novel (Pre-Training Assessment) and trained (Post-Training Assessment) audio-visual stimuli for each of the three analysed time windows of interest. Blue colour depicts actions paired with the non-verbal and red actions paired with the verbal referent (** $p < 0.01$, * $p < 0.05$ - green highlights significance from baseline, black indicates significance between conditions).*

4.4 Discussion

The aim of the present study was to investigate the functional involvement of action perception and language systems. We measured mu rhythm modulation triggered by linguistic and non-linguistic material in early toddlerhood at two points in time. To date no study has examined language induced motor resonance by directly comparing mu rhythm modulation before and after training to the same auditory material. The following main findings from our research efforts of this novel pre- post EEG assessment observational training study can be summarized as follows: as expected, no difference in mu rhythm was found between the two novel auditory referents at the initial EEG assessment prior to training. Post-training, mu desynchronization was only evident for the language referent. Topographical differences were observed between the familiar verb (left) and the trained verb post training (right). Motor activation triggered

by the audio-visual stimulus material also differed as a function of trained referent.

The motor activation from the audio-visual stimulus material indicated an overall decrease in motor activation from pre-training to post-training assessment. We also observed a reduction in activation between the two time points. Initially, all three time windows displayed mu desynchronisation independently for the actions observed (novel vs. familiar). Post-training, motor activation became more targeted and decreased to a mu supression within the third time window (the epoch representing action execution). This was only evident for actions trained with the verbal referent. No activation was observed in any of the time windows post training for actions trained with the non-verbal referent. A possible explanation could be that with increased experience with the action from training, toddlers started to contextualise the action, embedding it in the wider action context. The action learned in the language context, now became more meaningful, predictable and/or interpretable. Thus, the presence of a verbal label increased the familiarity of the action relative to the presence of the sound. First the action was primed with the verbal or nonverbal action referent (time window 1), this information was processed leading to action recognition of the action in the second half of the presented audio-visual material. Priming via the non verbal referent did not show the same activation. This hypothesis is in line with works that demonstrated significant mu desynchronization in response to familiar actions (e.g., Buccino et al, 2005; Calvo-Merino et al., 2005; 2006; van Elk et al., 2008).

How can the parallels in activation between the two audio-visual novel actions and the familiar action at initial assessment be explained? Evidence of similar patterns of activation can be found in literature examining practice effects. In a review, Kelly and Garavan (2005) describe the effects of neural changes with practice in the primary motor and sensory cortex. The authors argue that motor learning occurs in phases. At

first, performace is fast, typically occuring over early trials (associated with increased attention). This phase is then followed by a consolidation period. The final phase is characterized by slow learning, which strengthens the link between sensory or motor units. This final phase will eventually lead to a gradual increase in performance and is associated with increase in motor activation (see Kelly & Garavan, 2005 for an extended review). Similar findings are present in an adult fMRI study on the effect of a single practice session of hand actions (Voght et al., 2007). The authors found stronger premotor activation for observed non-practiced actions (guitar cords) when compared to practiced actions. This indicates a strong role for the mirror neuron system during the early stages of imitation.

This finding contradicts theories of motor activation solely dependent on the observer's motor repertoire. For instance Calvo-Merino and colleagues (2006) investigated the neural bases of motor activation induced by action observation by studying expert dancers. In their study, adult dancers showed greater activation in areas associated with the mirror neuron system for dance genres in which they were trained. Despite equal visual familiarity with particular dance moves due to conjoint training, motor activation was greater for moves that belonged to the dancers' own gender class and thus lay within the dancers motor repertoire. In Cannon and colleagues' (2014) study, adults who had acquired action experience prior to testing also showed higher levels of motor activation. Similar findings related to the mu rhythm in 14-16-month-olds comes from van Elk et al. (2008), where dependence on the ability to produce an action is documented. The paradigm of the present study did not allow for active experience and thus is specifically based on action observation. If the level of mu suppression were exclusively dependent on the ability to physically perform the observed actions, our results would yield no differences between the two trained

conditions. Our findings thus indicate that a) *changes in mu rhythm occur after a one-week observational experience with a novel action but only if* b) *the action was accompanied by language.*

This does not, however, explain why we observe activation across the three time windows similar to that of the familiar action at pre-training. One possibility could be that the familiar action were performed in novel context due to the use of the unusual and novel sand-like medium. This may have lead the toddlers to re-evaluate an already familiar action. As we did not assess motor activation to the familiar action a second time (other than in the auditory domain), we can only speculate that with more perceptual experience, toddlers would display similar neural patterns as that of the auditory only data. Furthermore it is to note that all three performed actions do convey sub-actions already in the toddler's motor repertoire (pincer grip and pointing with the index finger) meaning that the actions used were not exclusively novel as those used by Gerson et al. (2015). Consequently, it could be argued that first, in the new context, toddlers initially recognized these hand actions, however they ignored the action effector. This would support Buccino et al. (2004) who found activation of motor areas when observing actions performed by other species (e.g., dog barking). In the event of observed unfamiliar actions, one can decompose these actions into familiar components via motor resonance (Buccino et al., 2004).

De Klerk and colleagues (2014) postulated mere associative learning mechanisms in their study with 7- and 9-month olds. The findings of the present study suggest that toddlers have learned, or, to some extent, were able to represent the meaning of the learned action verb post-training. Not only do we find no evidence for mu desynchronization at time point 1 for both auditory referents, post-training assessment revealed that only the verbal-referent induced motor activation. This finding

is to some extent in agreement with the literature on language induced motor activation (e.g., van Elk et al., 2010 on language-induced motor resonance). However, here hemisperic differences were found between the familiar (left lateralized) and trained verbal-referent (right lateralized). There is one distinct difference that separates the familiar action verb squeezing from the trained action verb yokking. The familiar action referent is within the toddlers motor repertoire, whereas the trained verb is outside of the toddler's motor ability. Voght and colleagues (2007) found differences between the guitarists and non-guitarists in laterality during motor preparation and execution. More specifically the authors found the left hemishere to reflect more elaborate representations of the chord postures in the mirror neuron area. Nevertheless, the right hemisphere has indeed a functional role in comprehending language and more specifically word meaning (Federmeier, Wlotko, & Meyer, 2008). Rüschemeyer and colleagues (2007) found motor activation of more abstract verbs (e.g., to comprehend) to be more right laterized when compared to simple verbs (e.g., to grasp) (Rüschemeyer, Brass, & Frederici, 2007). The present study may therefore contribute to the understanding of interactive language processing mechanisms between the two cerebral hemispheres. Support for this notion comes from previous literature on active and passive training. Infants in Paulus et al.'s (2012) study showed a stronger mu desynchronization when listening to an action-related sound that the infants produced themselves following training. Familiar (trained) non-action-related or novel sounds could not induce such activation. However, it is to note that in the study by Paulus and colleagues (2012), the equally familiar sound from training lacked a visual referent. Infants in Gerson and colleagues (2015) study experienced active as well as observational training upon objects, which produced distinct sounds. Greater mu suppression was found to the sound associated with the action that the infants

performed themselves when compared to sounds from the passively observed action (Gerson, Bekkering, & Hunnius, 2015). The present study, however, found mu suppression to the trained verb without explicit motor experience. Therefore, motor experience does not explain the pattern of mu desynchronization observed over central sites in toddlers. Motor resonance displayed weaker activation to verbal stimuli when compared to viewing the videos in the present study. Fargier and colleagues (2012) observed a similar pattern of mu suppression, which can be interpreted as a form of reinstatement of brain activity. It can be speculated that with a prolonged training period and active experience language induced motor activation would likely parallel that of visual stimuli. This alteration would increase interpretability, and bring evidence if and how active learning impacts semantic formation. Due to the complexity of our study design, we have not included any control-conditions. For instance, it is an open question how motor activation would unfold in a group of 2-year-olds also assessed at two time points but in complete absence of home training. Also, how does language and action performance related to motor resonance? Further investigation is necessary in order for more in-depth conclusions to be made.

A difficult area in EEG analysis, is to find a good balance between analyses of a priori regions or time windows of interest and the need to take a reasonable account of the data as it presents itself. The initial selection of electrodes was driven a priori with channels covering the motor cortex as found in previous infant EEG literature (e.g., de Klerk et al., 2014). A standard approach to infant EEG is to also determine each dataset in terms of its topography (Hoehl & Wahl, 2012) resulting in channels that may overlap but it still maps on a priori components. A general issue in this field is the large variety of brain imaging methods used between adult and infant studies making it difficult to denote which channels on the sensor net used in our study, correspond with the reported

motor strip. In the present study, we further faced an issue in which we had to narrow down channels for the familiar control verb in order to directly compare the experimental conditions. As the familiar auditory material was more pronounced in terms of its latency and frequency band (6-9 Hz as opposed to the trained verbal referent which was concentrated around 8 Hz), we may have not been able to reflect the effect appropriately. Research must concentrate to solve these methodological issues.

4.5 Conclusion

This study is the first to investigate language induced motor resonance by directly comparing mu rhythm modulation at two time points. Verb processing distinctly modulated motor activation in early toddlerhood, indicating the close and functional involvement of action perception and language systems. Within the framework of embodied cognition, our findings highlight the importance of language during simultaneous representation of action performance. Semantic representation of verbs was hereby not strictly dependent on meaning-specific brain regions. Instead, auditory verbal information during the action learning process, also lead to the formation of synaptic connectivity between neurons in specific motor and language areas (Hauk et al., 2004; Pulvermüller, 2005; Fischer & Zwaan, 2008). This indicates that only when a linguistic description of an action is understood, motor-related regions become active. How active learning experiences may further reorganize neural structures during development and how this may shift topographically, still requires further investigation.

4.6 Authors Contributions

K.K., and L.F. conceptualized this study and were involved in the methodology and data acquisition. K.K. undertook data editing of the acquired data and K.K, and V.M.R.

performed the formal analysis. The original draft was written by K.K. who also provided visualizations. K.K, L.F., and V.M.R. were involved in review and editing and resources were kindly provided by D.M.

Chapter 5: General Discussion

5.1 Summary

Research on semantic processes generated by electrophysiological measures for both language and action domain separately, provides largely robust but also confounded evidence for the link between these two cognitive domains. How both domains knit together is still debatable in developmental research literature. Within the field of developmental cognitive neuroscience, measures for studying semantics are also not well defined. How the meaning, or semantics, of action symbols emerge and what role language plays in this process, as well as what measures can be applied to the assessment of semantics, has been the subject of the three experiments conducted as part of this doctoral thesis. The main findings are summarized here.

The first study explored the relation between language performance at 9 and 18 months and the ability to detect and interpret human actions at 9 months of age. The results of study one, with 28 monolingual Swedish infants, revealed that only children within a high proficiency language group at both time points successfully differentiated between the two action outcomes as indexed by the presence of the N400 ERP component for the incongruent action condition. Thus, the ability to judge whether an observed non-communicative human action is to be expected or not, on the semantic level, is closely related to language performance. This study replicated previous findings (Reid et al., 2009) using the same stimulus material and further, with the addition of the language measure, adds to the general understanding of mechanisms involved in the processing of sequential events that are shared between two conceptually different cognitive domains. A second ERP component examined in this

study was the Negative central (Nc) component, which is related to attention mechanisms. We observed Nc differences within the present dataset, which was congruent with prior findings by Reid et al. (2009). At the overall group level, 9-month-olds allocated more attentional resources when processing the expected actions. When the group was subsequently split according to language proficiency, however, neither of the language groups were significantly different in terms of their Nc component between the familiar and unfamiliar action outcomes. These results may suggest that all sub-groups are equally capable of detecting the familiar action outcomes from the unfamiliar ones and further allocate more attention to the more familiar expected action endings. In comparison with behavioural results, the ERP approach in this first study provides a basis for a stronger conclusion in terms of the specific cognitive activities related to processing action sequences. However, rather than proposing that relations exist between the action and language domain, the first study allows for the assessment of a specific cognitive mechanism in the form of the N400. As the N400 is involved specifically during semantic tasks, this cognitive process must be involved in the processing of actions. Therefore, it is possible to conclude that the semantic processes that are engaged during action observation are linked to language proficiency. We demonstrated that some basic cognitive mechanisms involved in the processing of sequential events are shared between these conceptually different cognitive domains. These mechanisms may even directly contribute to the acquisition of verb knowledge. Arguably the ability to parse actions in a meaningful way (e.g., Baldwin, et al., 2001; Sommerville & Woodward, 2005) may aid the ability to map these actions onto referents and generalize across action contexts.

The second study investigated the trajectory of new semantic knowledge formation in twenty 10 to 11-month-old infants via a highly social communicative live

learning situation in which an experimenter introduced and labelled two novel objects. Semantic processes were measured in a subsequent retention test shortly after learning, and with a 24-hour delay. The results suggest that new information transforms to a more long-term representation over time, as evidenced via the N400 ERP component for incongruent word-object pairs only in the subsequent retention test 24-hour post learning. In addition to the N400 ERP, we focused on Event-related Oscillations (ERO). More specifically, we examined upper alpha activity in early development. In adults, this upper alpha range reportedly activates the semantic knowledge system (Klimesch, 2012). Upper alpha band suppression for congruent word-object pairs and a correlation with the N400 ERP component was also evident in the delayed retention test. This suggests that this specific upper alpha frequency (ERO) may well be a representative neural marker for recruitment of the knowledge system also in early development. This implies that both correlates may index related cognitive mechanisms, namely the processing of semantics. In addition, we examined differences in attentional mechanisms in the immediate and delayed retention phase. Similar to the observed N400 ERP and upper alpha ERO effect, the Nc, associated with allocation of attention, was greater for incongruous word-object pairs only 24-hours later. Given these consistent findings, this study highlights the importance of offline consolidation and proposes that semantic representation of newly acquired words in memory is most likely a function of time. This study is the first study to demonstrate and to relate infant modulation of posterior upper alpha (9-10 Hz) and N400 as two representative markers for semantic processing and recruitment of the knowledge system. Thus, our findings contribute to our understanding of upper alpha band suppression and its involvement in integration and retrieving knowledge from the knowledge system. Further, it supports the notion of the existence of distinct sub-bands, similar to those reported in adult

literature (Klimesch, 2012). These results indicate that infants can learn semantic information rather than rely on associative learning as an underlying mechanism for social processing. In order to fully enable these processes to come to light, however, one needs to consider that consolidation of new information happens over time.

The final study directly compared modulation of the mu rhythm in twenty 24-month-old children and related suppression of mu in response to action-related auditory and visual verbal material in contrast to non-verbal material, before and after training. As expected, results of this final experiment revealed no significant differences between the two novel auditory or visual action stimuli at pre-training assessment. Post-training, we observed significant differences in mu desynchronization in response to the newly acquired auditory verbal-referent, however, mu activation did not parallel that of a familiar control action verb (an action which was present in the toddlers' motor repertoire). Visually induced motor resonance to the novel and trained hand actions also differed as a function of training. While trained action videos, primed by the non-verbal referent, did not show mu desynchronization, audio-visual stimulus material triggered by the trained verbal referent lead to significant motor activation from baseline during action observation. These results suggest that verb processing distinctly modulates motor activation in early toddlerhood, indicating the presence of a close and functional involvement of action perception and language systems. Within the framework of embodied cognition (e.g., Fischer & Zwaan, 2008; Gallese, 2008; Glenberg & Gallese, 2012; Pulvermüller, 2005), these findings highlight the importance of language during simultaneous representation of action performance. Semantic representation of verbs is therefore not strictly dependent on meaning-specific brain regions. Instead, auditory verbal information during the action learning process, can also lead to the formation of synaptic connectivity between neurons in specific motor and language areas. This

indicates that only when a linguistic description of an action is understood, motor-related regions become active.

5.2 Theoretical implications

The present studies, when combined, have provided an array of new knowledge related to how action and language are linked to one another, and how semantics plays a role in the interactions between these two cognitive domains. Despite this, the ontogeny of action and language development and especially the possible relation between these two cognitive domains is still not well understood. A general question in developmental research and in developmental cognitive neuroscience is how the meaning or semantics of action symbols emerge and the role of language within this process. For instance, language development may, in part, stem from the ability to make sense of others' actions. Learned expectations about actions and action outcomes, whether shaped by observation or whether self-experienced, can act as a scaffold for later language development. The position that was adopted in this thesis is that, in order for language and action understanding to evolve, anything newly learned will inevitably need to be stored in memory and translated into meaning. Equally, to be able to act and react to the external and internal incoming and outgoing information streams, these stored meaningful memories require access. Thus, we explored the notion that both cognitive domains have one common ground, namely their reliance to consult their semantic meaning centre as proposed by Rüschemeyer et al. (2009). We designed three targeted studies aimed to reflect different stages in early development to monitor the processes of accessing semantic systems, directly as they emerge, during early development. Thus, our main emphasis was placed on semantics within the action and the language domain. The general results of the present thesis can be summed up as followed:

Semantics is interwoven in both action and in language processing and in their interactions. The section below will provide implications from our findings to the current understanding of action and language semantics.

5.2.1 Bidirectional account of Action and Language

Evidence that aspects of social understanding, such as the ability to interpret eye gaze (Brooks & Meltzoff, 2005), gestures (Butterworth and Morissette, 1996, Rowe & Goldwin-Meadow, 2009a, 2009b; Tomasello et al., 2007) including pointing (Camaioni et al., 1991) or understand joint attention (e.g., Tomasello & Farrar, 1986) predicts later language proficiency (e.g., Baldwin, 1993; Iverson & Goldin-Measow, 2005; Kuhl et al., 2003), is already evident in the developmental literature. Whether language performance is related to the ability to detect and interpret human actions at 9 months of age, however, remained unaddressed. We proposed that the origins of processing and understanding hierarchical representations are to be found in infancy and that the hierarchical nature of actions can, to some extent, help facilitate action predictions by means of semantic rules. Our results suggest that early language competence is aligned with semantic processing abilities in the action domain during infancy. Further, we aimed to highlight the noticeable parallels in the organization of actions and the structure of language. We found evidence for this notion with effects occurring earlier in ontogeny when compared with other known social competencies that are predictive of language (e.g., Brooks & Meltzoff, 2005; 2008). The current study therefore links with prior language development research and provides a different pathway in the attempt to understand the development of language by investigating non-communicative human actions and by using neurophysiological measures of action semantics reflected in the N400 ERP component. We suggest that language proficiency

is aligned with semantic processing abilities during development. To date, literature could not comment on the underlying structure of cognitive processes as the prior literature was based on inferences from behavioural results. We highlight the parallels in the organization of actions and language, and demonstrate that some basic cognitive mechanisms must be involved in the processing of sequential events that are shared between the two conceptually different cognitive domains of action and language. Moreover, these basic cognitive mechanisms seem to already coexist early in development. Consequently, we raise an intriguing concept that understanding actions during early development may, to some extent, evolve at the same time as language acquisition or, perhaps, act as a mediator for language proficiency. Finally, we provide important information related to development by examining data via an individual differences approach, rather than making more general inferences from overall group level effects.

5.2.2 Neurophysiological Measures of Semantics within Context: The Indirect Route

5.2.2a Upper Alpha

Semantic orientation enables individuals to selectively access stored information such as sensory information and 'higher order information', including language, mathematics, and geography (Klimesch, 2012). This can be reflected in upper alpha band desynchronization for semantically integrated information. Thus, desynchronization of activation in the "upper" alpha range has been associated with accessing a semantic knowledge system in adults, which could play a role in learning (Klimesch, 2012). Prior to our investigations, the involvement of "upper" alpha band desynchronization as a potential and *complementary* neurophysiological marker to the N400 ERP component in early development was not explored. We hypothesized that we

would find upper alpha band desynchronization (9-10 Hz) in early development, similar to that of adults (Klimesch, 2012). This leads to the suggestion that the specific manifestation of alpha may well represent a signature of constructing the knowledge system during infancy. We looked at possible relations between these two correlates to help answer the question of whether both neural markers index the same cognitive mechanism, namely the processing of semantics. Our findings not only indicate that the N400 can serve as a useful tool to test semantic memory, but also that the suppression in the upper alpha band may well reflect the ability to learn and remember what words refers to in early development. The presence of alpha band suppression suggests that this specific manifestation of alpha is a signature of constructing the knowledge system at 10- to 11-months, further suggesting the possible existence of multiple sub-bands in the alpha range. Upper alpha desynchronization and the N400 are to some extent related to one another, with infants who displayed an N400 for incongruent word-object pairs in the delayed retention test also likely to show greater alpha suppression for congruent word-object pairs when compared to incongruent pairs. This points towards the notion that both neural markers may well represent related cognitive mechanisms.

5.2.2b Validity for Language and Action N400

Results from infant studies using the N400 neural correlate as a tool to investigate semantic processing are currently inconsistent. The first two experiments have investigated the N400 semantic measure from two distinct directions. In study one, we looked at the action N400 by violating 9-month-olds expectation of a familiar action outcome. In the second study, we primed 10-11 month olds with a familiarized word and subsequently violated infants' expectation by presenting the wrong object referent. In both of these studies, the N400 morphology was observed over posterior sites in the typical time window and with the typical morphology (N400 in the incongruent

condition only) peaking at around 700 ms post stimulus presentation (see Reid et al., 2009, for similar results). The right topographical distribution for the observed N400 of the presented research findings stands further in accordance with the reported right bias for the N400, which is further parietally maximal within adult literature (see Kutas & Federmeier, 2011). Thus, for both cognitive domains, the N400 was found to be a sensitive and reliable indicator of the semantic relationship between the target and context. In other words, the N400 component is prone to elicit more negativity for infant participants when the target does not fit the semantic representation of the context in which the target occurred.

5.2.3 The Role of Infant Directed Speech for Semantic Formation

Infant Directed Speech has highly attention inducing properties, which promote attention to speech, word-object mapping and word recognition of newly acquired words (Ma et al., 2011; Singh et al., 2009; Zangl et al., 2007). Singh et al. (2009), for instance, found IDS to be crucial during training of new words in 7.5 month olds with effects on implementation into memory observed 24-hours later. Further, if the labels were delivered in an ID manner during training, recall was not affected by the fashion via which the words were presented at test (ID vs. AD). Contradicting evidence comes from Parise and Csibra (2012) who could only find evidence for the N400 component to the incongruent word-object pairs only if their mother, but not an experimenter, introduced those objects in a live paradigm. They consequently concluded that word-to-object priming occurs referentially in 9-month-olds only if the referential nature of words is conveyed through the mother. Unlike Parise et al. (2008), the Nc effect observed by Kopp and Lindenberger (2011) serves as evidence that 9-month-olds are able to distinguish between new and old labels irrespective of the degree of joint

attention during the initial familiarization phase. Although these studies utilized different measures to study early word recognition, what becomes evident is that inconsistencies in current literature on the role of IDS are still prevalent in the field. Singh et al. (2009) pointed out that young language learners will, with time, become more attuned to ADS. At what time in development the influence of IDS reduces in importance, however, requires further investigation. The implications that we can draw from the obtained results of the second experiment are as follows: Both novel objects were presented and labelled in a highly social live way and were accompanied by other social referential cues such as eye contact between the experimenter and infant for approximately 6 minutes. During the recall, the familiarized words were introduced by a neutral tone of voice and in addition by a second female voice. Thus, at recall test, the voice priming the objects on the screen was not the initial experimenter from the learning phase. It should be noted that the infants of the present study were of slightly older age when compared to the infants from Parise and Csibra's study (2012) (10 to11 as opposed to 9 months) thus we could speculate that already at this slightly older stage in development, IDS reduces in importance, leading to successful activation of the object features associated with familiarized objects primed by a stranger in a neutral manner. In experiment three, the female voice during training and test was the same, however in both instances the speaker did not unitize IDS. The results of study three indicated that 24-month olds could form stable semantic representations, as indexed by distinct mu desynchronization of the language referent, from pre- to post-training assessment. It therefore seems that, by the second postnatal year, toddlers are becoming increasingly attuned to ADS, which leads to successful verb learning despite reduced social cueing during learning and recall.

5.2.4 The Role of Imitation

Early imitation plays an important role in providing infants with a general understanding of how it feels to be someone else (Meltzoff & Decety, 2003). In addition, within the mirror neuron account, (e.g., Pineda, 2005; van Elk et al., 2008; Reid et al., 2011) it is claimed that observed actions could only be fully understood if they can be mapped onto one's own motor representations (e.g., Nyström et al., 2008). Although we support the importance of early imitation, the findings from our third study on motor resonance indicate that the role of imitation may also decrease in importance over time as indexed by modulation of mu from observing actions during training. Thus, similar to the example outlined above, of the influence of IDS for word processing, these findings imply that sensorimotor experiences become increasingly independent for the formation of language and action systems (Piaget, 1952). More recently Gampe et al. (2016) highlighted the importance of toddlers' imitation proficiency for verb learning success as this may increase the likelihood to learn a novel label. However, this effect was only true for the younger children of their sample (24-30-month). Older children (36-month) did not seem to rely on action reproduction during the verb learning process. Consequently, it can be concluded that action reproduction, or imitation of the actions of others, as well as IDS, decrease in importance with development and that instructions via language become more meaningful.

5.2.5 The Role of Joint Attention

The issues raised for IDS and imitation should also be examined for joint attention mechanisms. Parise et al. (2008) found evidence for enhanced object processing in 5-month-olds, as indicated by a more negative Nc component to objects in which eye

contact took place in a pre-test phase**.** The authors suggested that seeing someone directing attention to an environmental object induces a shift of attention in the infant, resulting in the alignment of both subjects' attention on to the same object. This implies that infants profit from joint attention when processing information in their surrounding environment. That joint attention is still central in adulthood during object processing can still be speculated. Lachat and colleagues (2012) for instance, found support for the importance of joint attention. In their study, increased alpha band desynchronization was evident when two adult participants focused on the same object subsequent to shared eye contact. This effect, utilizing ERO's was also found in infancy. Hoehl et al. (2014) for instance, found increases in brain activity in 9-month-olds (as indexed by lower alpha band suppression) to objects after an episode of joint attention. Michel et al. (2015) report similar findings during the perception of faces with object directed eye gaze. Although we didn't account for alpha in the lower range in the present thesis, it still highlights the importance of social aspects for object processing in development, arguably across the lifespan.

5.2.6 The Role of Consolidation for Semantic Formation

There are conflicts in the current literature on semantics and memory processes for newly acquired words, as a function of time. Whereas some studies report the disappearance of semantic priming after 24-hours (Friedrich & Friederici, 2011) more recent findings support offline consolidation and argue for its importance for the formation of semantic knowledge (Brown et al., 2012, Friedrich et al., 2014; Henderson et al., 2012; 2013). Consistent across all significant results of the second experiment is the evidence for semantic and memory processes. These were evident in the delayed retention test. We drew the conclusion that immediately after the learning experience,

some lower level associations most likely took place as the 10-11-month-old infants could not process information semantically and thus did not distinguish between correct and incorrect word-object pairs. By the time of the second visit to the laboratory 24-hour later, infants had most likely undergone a process of offline consolidation. This can reliably be shown via three neural correlates measured in this study (Nc, N400 & upper alpha suppression) leading infants' to successfully make inferences on the basis of semantic memory and attentional mechanisms. Our findings are in good agreement with recent findings from Friedrich and colleagues (2014) that also found no evidence for the N400 component during training, which indicates that lexical semantic knowledge is not acquired online during the training session of the two age groups tested (9 and 16 months). We can thus contribute to the given literature, that also within the age of 10- to 11-months, offline consolidation processes are necessary for semantic formation.

5.2.7 Neurophysiological Measure of Context Independent Semantics: The Direct Route of Mu Desynchronization.

5.2.7a The Role of Active Experience and Passive Experience

The act of observing others clearly influences a large proportion of an infant's action development. With increased proficiency in motor skills, further learning in motor specific (Daum et al., 2011) and non-motor specific domains (e.g., Brandone, 2015; Libertus & Violi, 2016; Oudgeneog-Paz et al., 2012; 2013; Pineda, 2005; Reid et al., 2011; Soska et al., 2010; Walle & Campos, 2014) will take place. The interplay between perception and action may therefore form the basis for stable associations between motor acts and sensory consequences and subsequently lead to complex cognitive functions, such as understanding others' goals and intentions (Hunnius & Bekkering, 2014). Even though this may imply the importance of both perception and

action, a motor resonance perspective puts more emphasis on action experience. Thus, in order to generate motor resonance, actions need to be well established in the observer's mind (e.g., Buccino et al., 2005; Cannon et al., 2014; van Elk et al., 2008; Calvo-Merino et al., 2006; de Klerk et al., 2014, Gerson et al., 2015; Nyström, et al., 2011). However, literature also reports evidence for the notion that observed actions may well be understood while lying outside one's own action capacities (e.g., Cross et al., 2012; Grossmann et al., 2014; Fischer & Zwaan, 2008). The paradigm of the third experimental chapter did not allow for active experience and thus is specifically based on action perception. Mu suppression triggered by the audio-visual manual actions was evident and differed distinctively as a function of training. Therefore, our results support the notion that at 24 months, learning about actions is not exclusively dependent on the ability to physically perform the observed actions. Observational experience still represents a stable source of information for action processing (Hunnius & Bekkering, 2014).

Observational training in this thesis not only had an impact on motor activation induced by visual material, it also proved to be influential for verb processing in our sample of 24-month-olds. Modulation of motor resonance induced by linguistic material has not been studied to the same extent as that of action perception in early development. Our findings show mu desynchronization to be distinctly different depending on the type of referent used during training (language vs. non-action related sound). Our findings therefore, further imply a close functional involvement of both action and language systems, given that both verbs and action perception can be reflected by the same neural correlate within the motor area. A further indicator that activation of motor areas is indeed specific to mechanisms of action and language processing can be seen by the effects observed for the non-verbal referents presented

alone. They did not lead to similar effects at post-test assessment when contrasted with the language familiar and trained referents. Thus, suppression of mu in response to action-related auditory verbal material can be distinguished from that of auditory non-verbal material, similar to the rattle in Paulus et al.'s, (2012) study. This result suggests that toddlers have learned, or, to some extent, were able to represent the meaning of the learned action verb post training rather than forming mere associations as postulated by de Klerk and colleagues (2014).

Overall, motor resonance displayed weaker activation to verbal stimuli when compared to visually presented actions; a pattern that parallels the findings of Fargier et al. (2012). We could speculate that prolonged training exposure, together with adding active experience, would lead to a more aligned motor activation induced by auditory when compared to visual action stimuli. Our results suggest that there are early links between motor and language systems that may operate at different levels of information processing.

5.2.7b Hemispheric Specialization

Mechanisms for intercerebral interaction for language and action in early development are also unknown. Our results of study three, indicate that the familiarized action verb at post-test assessment did not have the same EEG topographies as the putatively more embodied action verb. We therefore raise the possiblity that the processing of verbs currently outside ones' motor repertoire (similar to more abstract verbs in Rüschemeyer et al's., 2007 study) could show a different topographical distribution when compared to verbs within the toddlers' motor repertoire, such as "squeezing". The present study may therefore contribute to the understanding of interactive language processing mechanisms between the two cerebral hemispheres during development.

5.3 Limitations and directions for future research

The questions that were asked in the framework of this thesis were a) how does semantics emerge and how is semantic information stored and accessed by the action and language domains, b) what is the dynamic between these two conceptually different cognitive domains, c) do current measures of semantics index the same construct or different aspects of the same construct, and d) are there possible alternative measures that could be useful for measuring semantics in future research? Our main emphasis was placed on semantics within the action and the language domains. The general results of the present thesis can be summed up as followed: *"Semantics is required for processing actions and language successfully"*. However, there are still outstanding issues that we cannot conclude from our research efforts.

5.3.1 Specific limitations of this thesis

5.3.1a Theta Synchronization and Alpha Desynchronization

As part of our secondary hypothesis of experimental chapter two, we aimed to investigate theta band synchronisation as a possible indicator for "encoding" (Begus et al., 2015; 2016; Klimesch et al., 1996; Saby & Marshall, 2012). The violation of expectancy paradigm was designed to primarily account for the alpha frequency band and therefore was not well fitted to the investigation of theta. Not surprisingly, we found no clear evidence for active encoding as indexed by the absence of theta band synchronization. This was possibly because encoding of information had already taken place during the interactive learning phase. The investigation of infants' theta activation during the learning phase should therefore represent an intriguing possibility for further investigation. Evidence of theta during learning could serve as an indicator for successful semantic recruitment. Future research could therefore investigate this notion

by looking at theta activation within the learning context itself. In a related sense, with respect to upper alpha desynchronization, an interesting line of research could be to investigate neural responses more directly within the word prime before any visual cues appear. It is possible that successful recruitment of the knowledge system could already be activated during the build up of the mental representation triggered by language similarly to language induced motor resonance.

5.4.1b Complexity of Study Design

The design of the third study performed as part of this thesis was highly complex and involved too many variables. Consequently, many of the conditions did not enter the analyses simply because toddlers did not attend to all of the blocks of stimuli. Although we believe that the research questions that could have been answered, but were not assessed due to insufficient data, (e.g. N400 within the 4th Block at post-training assessment) would have added value. We could have perhaps simplified the procedure. For instance, we could have collected more artifact-free trials within the pure visual block (Block 2) by not presenting Block 3 (audio-visual) at all. The rationale for this position is that we cannot say with confidence that during the simultaneous presentation of the action and the referent at pre-and post-test, some learning processes were already (for pre-training) or still (for post-training) on-going. Further, presenting the familiar actions in the mixed modality during the pre-training phase (Block 4) was not carefully enough considered. Again, the auditory presentation could have served as a prime for subsequent action recognition (e.g., hearing the verb "squeezing" before presenting the visual equivalent of the squeezing action). This priming could be bidirectional, with the consequence that the visual and audio-visual familiar actions could help activate semantic knowledge as indexed by mu desynchronization within the auditory verb

presentation. Consequently, for an exploratory study such as this, the modalities should be kept in separate blocks rather than being presented within one self-contained block.

5.3.2 General limitations of the field

5.3.2a Validity of Language Measures

The general view across the literature points towards a weakness of parental reports in terms of determining comprehension capacities of their child (e.g., Friedrich and Friederici, 2010; Mills et al., 2005; Rämä et al., 2013, Torkildsen et al., 2006, Torkildsen et al., 2008; Torkildsen et al., 2009). We divided our sample at the 9 months comprehension level as well as at the 18 months language production level. Children were found to be equally likely to remain within the same language proficiency group as they were to move up or down in their language proficiency between the 9 months comprehension and the 18 months production assessment. It is therefore possible that the parental measures were creating effects because of their lack of robustness. Furthermore, at the time point of our second language assessment at 18 months, toddlers' productive vocabulary is still fluctuating. Based on these observations, we propose that results based on the comprehension measure at 9 months, should be taken with caution. Specifically, measures of language capacities are thought to become increasingly more consistent later in development (Bornstein & Putnick, 2012). To account for this, directions for future research could, for instance, implement a more longitudinal approach of language development and could be additionally followed at different points across development.

Furthermore, based on the results of our other studies, utilizing the N400 *and* upper alpha as primary tools could help to further determine the associations between action and language processing during early development.

5.3.2b Were the observed actions truly novel?

The two novel actions used in study three conveyed sub-actions that are already within a toddler's motor repertoire at two years of age (pincer grip and pointing with the index finger). As a result, and this is a wider issue within the field, we cannot claim that the presented actions were exclusively novel to the children at pre-training assessment. This could explain why we found motor activation that is similar to the familiar action at initial assessment. On the other hand, we nevertheless observed differences in motor activation as a function of training. Future research should consider the inclusion of the same familiar action also in the form of visual presentations at post-test (in the present study only in the form of the auditory modality). This would give stronger evidence that the modulation observed in this experiment is to be differentiated in comparison to the familiar control variable. Thus we would expect motor activation triggered by actions within the motor repertoire, to remain stable across both time points.

5.3.2c Consistency of channel inclusion for analysis

We found it difficult to denote which channels on the sensor net correspond with the reported motor strip from adult studies using other brain imaging methods. It appears that across adult and developmental literature, the channels chosen to enter the final analysis differ dramatically. This could be a function of paradigm and type of EEG used. In study three, we found evidence for a differing hemispheric activation between the familiar and trained verb post-training. The latency and frequency band was more pronounced for the familiar verb starting at the onset of presentation. Further motor activation to the familiar control verb displayed a wider frequency band (6-9 Hz) when compared to the trained verb (8 Hz). However, in order to directly compare the experimental conditions, we chose the time window of 600-800 ms post stimulus presentation at 8 Hz for 5 left central (to represent the familiar verb) and 5 opposite

right-central (to represent the trained verb) electrodes. As a consequence of these differences in topography, we neglected the wider frequency band evident in the familiar verb condition. In was in this manner that we could directly compare the two variables. Further research should be pursued to solve these issues of a methodological nature.

5.3.3 Future directions

5.3.3a The influence of action verbs within the learning context

The active social live learning phase enabled the infant and experimenter to jointly engage and explore the objects in the second experiment. The experimenter would, for instance, show the infant how the object could be held, shaken or rolled. This incidental use of action related words during learning could have enhanced learning and facilitated recognition due to its salience with respect to semantics. Infants may have allocated more meaning to the objects as a function of what the objects are used for. Findings from Hauf et al. (2007) further suggest that shared experience with the same object facilitates interest towards the object itself when compared to sole object exploration by the infant. Also, infants were able to actively perform the same actions upon the novel objects and imitate the experimenter. This additional exploration may have further bootstrapped word-object mapping. Thus, we cannot make strong conclusions as to whether the same effects could be if infants a) did not have the chance to hold the objects themselves and act upon them and b) if infants would form semantic representation in a less meaningful learning context (e.g., reduction of action words) while both adult and infant would still jointly interact and by keeping the utterances constant for object labelling during learning.

5.3.3b Motor Resonance – Semantic Expectancy Paradigm

Other intriguing findings deriving from motor resonance research, that are relevant for the ontogeny of language and action systems, report the possible involvement of mu rhythm and semantic processes (Uithol et al., 2011). Stapel and colleagues (2010) report greater mu desynchronization when 12-month-olds saw unusual actions such as bringing a cup to the ear in comparison to actions that complied with the infant's expectancy of how a cup should be used (e.g. cup lifted and brought towards the mouth). It is possible that semantic congruence affects the mirror neuron system by enhancing activity to incongruent actions. Future research should therefore investigate possible links between the N400 ERP component and the infants emerging mirror system in a semantic congruency task to establish whether these two indirect routes to semantic processes develop in tandem.

5.3.3c Additional concepts for future research

The next logical steps for future research would be to include active training and monitor motor activation as a function of active experience. Also, sound used in future research could be action related rather than a two-tone sequence. Including two novel sounds at post-test of the third study could be also followed up for the visual modality. In addition, the familiar visual action should also be present at post-training assessment. These considerations would allow for better interpretability, as it would clarify if active learning strengthens semantic formation as indexed by language induced motor resonance. Under this scenario, we would expect mu rhythm to parallel that of observed actions. Further, it would be intriguing to investigate whether mu would also be more supressed when viewing still images of the mid action state before and after active and/or passive training. We should not forget to relate all the above with language and action performance. Thus, assessing language skills and looking at individual

differences in performing the actions post-training could give valuable insights into individual differences within this body of research.

Finally, in study two, only two novel objects and two novel labels were used to simplify learning. Interesting research could derive from these finding and investigate generalization and categorization abilities by presenting more objects during learning. Furthermore, introducing a control variable to monitor how the N400 as well as upper alpha behave in complete congruency would also widen our knowledge about functional properties of these neural correlates.

5.4 Conclusion

The primary focus of this doctoral thesis was placed on semantics. Within this framework, we investigated how semantic information develops and how semantic representations are stored and accessed by two conceptually different, but closely linked, cognitive domains in early development. The results from the studies in this thesis allow us to draw the conclusion that for both the language and action domains, semantic knowledge is key. Furthermore, we argue that the semantics construct is shared between these two domains. We put forward a proposal that both language and action are closely related. It could be argued that the action domain in early development is to some extent overruling and crucial for language development. With increasing age, both domains may become equal in importance for the rapid progress of higher order cognitive skills. At some point, however, this dynamic may change entirely and it remains open if language becomes a more dominant cognitive construct later in life. We believe that we cannot answer any of these questions with confidence as a function of the data presented in this thesis. However, we hope to have identified some intriguing concepts in order to come one step closer to understanding the dynamic

interplay between action and language. The field of developmental cognitive neuroscience should be encouraged to utilize complementary methods, such as those outlined in this thesis, to further our knowledge on these issues.

References

Axelsson, E.L., Williams, S.E., & Horst, J.S. (2016). The effect of sleep on children's word retention and generalization, *Frontiers in human neuroscience*, *7*(1192).

Amoruso, L., Gelormini, C., Aboitiz, F., González, M.A., Manes, F., Cardona, J. F., & Ibanez, A. (2013). N400 ERPs for actions: building meaning in context. *Frontiers in human neuroscience*, *7*(57).

Bakker, M., Kaduk, K., Elsner, C., Juvrud, J., & Gredebäck, G. (2015). The neural basis of non-verbal communication – enhanced processing of perceived give-me gestures in 9-month-old girls. *Frontiers in Psychology, 6(59),* 1-6.

Baldwin, D.A. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology, 29*(5), 832-843.

Baldwin, D.A., Baird, J.A., Saylor, M.M., & Clark, M.A. (2001). Infants parse dynamic action. *Child Development*, *72*(3), 708–717.

Begus, K., Gliga, T., & Southgate, V. (2016). Infants' preferences for native speakers are associated with an expectation of information. *Proceedings of the National Academy of Science, 113*(44), 12397-12402.

Begus, K., Southgate, V., & Gliga, T. (2015). Neural mechanisms of infant learning: differences in frontal theta activity during object exploration modulate subsequent object recognition. *Biology Letters, 11*(5).

Bendy, M., & Caramazza, A. (2011). Perception, action, and word meaning in the human brain: the case from action verbs. *Annals of the New York Academy of Science, 1224*, 81-95.

Berglund, E., & Eriksson, M. (2000). Communicative development in Swedish children 16–28 months old: The Swedish Early Communicative Development Inventory-words and sentences. *Scandinavian Journal of Psychology, 41*, 133-144.

Bertenthal, B.I., Pinto, J. (1993). Dynamical constraints in the perception and production of human movements, In M. Gunnar & E. Thelen (Eds.), *The Minnesota Symposia on Child Psychology: Systems in Development* (pp. 209-239). Erlbaum, Hillsdale, N.J,

Bertenthal, B.I., Proffitt, D.R., Kramer, & S.J. (1987). Perception of biomechanical motions by infants: Implementation of various processing constraints. *Journal of Experimental Psychology: Human Perception and Performance, 13*(4), 577-585.

Blasi, A., Mercure, E., Lloyd-Fox, S., Thomson, A., Brammer, M., Sauter, D., Deeley, Q. Barker, G.J., Renvall, Deoni, S., Gasston, D., Williams, S.C.R., Johnson, M.H., Simmons, A., & Murphy, D.G.M. (2011. Early specialization for voice and emotion processing in the infant brain. *Current Biology, 21*(14), 1220-1224.

Borgström, K., von Koss Torkildsen, J., & Lindgren, M. (2017). Visual event-related potentials to novel objects predict rapid word learning ability in 20-month-olds. *Developmental Neuropsychology,* 1-16.

Bornstein, M.H., & Putnick, D.L. (2012). Stability of language in childhood: A multiage, multidomain, multimeasure, and multisource study. *Developmental Psychology, 48*, 477-491.

Brandone, A.C. (2015), Infants social and motor experience and the emerging understanding of intetnional actions. *Developmental Psychology, 51*(4), 512-523

Brooks, R. & Meltzoff, A.N. (2005). The development of gaze following and its relation to language. *Developmental Science, 8*(6), 535–543.

Brooks, R., & Meltzoff, A.N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal growth curve modelling study. *Journal of Child Language, 35*, 207-220.

Brooks, R., & Meltzoff, A.N. (2015). Connecting the dots from infancy to childhood: A longitudinal study connecting gaze following, language, and explicit theory of mind. *Journal of Experimental Child Psychology, 130*, 67–78.

Brown, H., Weighall, A., Henderson, L.M., & Gaskell, M.G. (2012). Enhanced recognition and recall of new words in 7- and 12-year-olds following a period of offline consolidation. *Journal of Experimental Child Psychology, 112*(1), 56-72.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action related sentences modulates the activity of the motor system: a combined TMS and behavioural study. *Cognition Brain Research, 24*, 355-363.

Butterworth, G.E., & Morissette, P. (1996). Onset of pointing and the acquisition of

language in infancy. *Journal of Reproductive and Infant Psychology*, 14, 219–231.

Calvo-Merino, B., Glaser, D.E., Grèzes, J., Passingham, R.E., & Haggard, P. (2005). Action observation and acquired motor skills: an FMRI study with expert dancers. *Cerebral Cortex, 15(8),* 1243-1249.

Calvo-Merino, B., Grèzes, J., Glaser, D.E., Passingham, R.E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology, 16(19),* 1905-1910.

Camaioni, L, Caselli, M.C., Longobardi, E., & Volterra, V. (1991). 'A parent report instrument for early language assessment'. *First Language*, *11,* 345-59.

Cannon, E.N., Yoo, KH., Vanderwert, RE., Ferrari, PF., Woodward, AL., & Fox, NA. (2014). Action Experience, more than observation, influences mu rhythm desynchronization. *PLoS ONE, 9(3):* e92002.

Caramazza, A., Anzellotti, S., Strnad, L., & Lingnau, A. (2014). Embodied cognition and mirror neurons: a critical assessment. *Annual Review of Neuroscience, 37,* 1-15.

Corballis, M.C. (2003). From hand to mouth: The gestural origins of language. In M. H. Christiansen & S. Kirby (Eds.), Language evolution (pp. 201-218). New York: Oxford University Press.

Cross, E.S., Liepeilt, R., de C. Hamilton, A.F., Parkinson, J., Ramsey, R., Stadler, W., & Prinz, W. (2011) Robotic actions preferentially engage the human action observation network. *Human Brain Mapping, 33*(9), 2238-2254.

Csibra, G., Davis, G., Spratling, M.W., Johnson, M.H. (2000). Gamma oscillations and object processing in the infant brain. *Science, 290*(5496), 1582-1585.

Cuevas, K., Cannon, E.N., Yoo, K., & Fox, N.A. (2014). The infant EEG mu rhythm: methodological considerations and best practices. *Developmental Review*, *34*(1), 26-43.

Daum, M.M., Prinz, W., & Aschersleben, G. (2008). Encoding the goal of an object directed but uncompleted reaching action in 6- and 9-month-old infants. *Developmental Science*, *11*(4), 607–619.

Daum, M.M., Prinz, W., & Aschersleben, G. (2011). Perception and production of object-related grasping. *Journal of Experimental Child Psychology, 108*(4), 810-818.

Daum, M.M., Ulber, J., & Gredebäck, G. (2013). The development of pointing comprehension in infancy: Effects of communicative verbal cues on covert shifts of attention. *Developmental Psychology, 49(*10), 1898-1908.

de Haan, M. (2007). Visual attention and recognition memory in infancy. In M. de Haan (Ed.), *Infant EEG and Event-Related Potentials* (pp. 101-143). Psychology Press: Hove and New York.

de Klerk, C.C.J.M., Johnson, M.H., Heyes, C.M., & Southgate, V. (2015). Baby Steps: investigating the development of perceptual-motor couplings in infancy, *Developmental Science, 18,* 270-280.

de Haan, M., & Nelson, C.A. (1997). Recognition of the mother's face by six-month-old infants: A neurobehavioral study. *Child Development, 68*, 187-210.

Domínguez-Martínez, E., Parise, E., Strandvall, T., & Reid, V.M. (2015). The fixation distance to the stimulus influences ERP quality: An EEG and eye tracking N400 study. *PLoS ONE, 10*(7).

Domínguez-Martínez, E. (2017). *Towards the improvement of event-related potential techniques in developmental populations:* (PhD thesis). Lancaster University, Lancaster, United Kingdom.

Eriksson, M., Westerlund, M., & Berglund, E. (2002). A screening version of the Swedish Communicative Development Inventories designed for use with 18-month-oldchildren. *Journal of Speech Language and Hearing Research, 45*, 948-960.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology, 73 (6)*, 2608-2611.

Fargier, R., Paulignan, Y., Boulenger, V., Monaghan, P., Reboul, A., & Nazir, T.A. (2012). Learning to associate novel words with motor actions: language induced-motor activity following short training. *Cortex, 48*(7), 888-899.

Farroni, T., Csibra, G., Simion, F., & Johnson, M.H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America, 99*(14), 9602-9605.

Federmeier, K.D., Wlotko, E.W., & Meyer, A.M., (2008). What's 'right' in language comprehension: event-related potentials reveal right hemisphere language capabilities. *Language and Linguistics Compass, 2*(1), 1-17.

Fenson, L., Dale, J.S., Reznick, D., Thal, E., Bates, J., Hartung, S., Pethick, J., & Reilly, J. (1993). MacArthur Communicative Inventories: User's guide and technical manual. San Diego: Singular Publishing.

Fenson, L., Dale, P.S., Reznick, J. S., Bates, E., Thal, D.J., & Pethick, S.J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5), 1-185.

Fischer, MH., & Zwaan, RA. (2008). Embodied language: a review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology, 61*, 825-850.

Fogassi, L., Ferrari, P.F. Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science, 308*(5722), 662-667.

Fourkas, AD., Bonavolonta, V., Avenanti, A., & Aglioti, SM. (2008). Kinesthetic imagery and tool-specifc modulation of corticospinal representations in expert tennis players. *Cerebral Cortex, 18(10),* 2382-2390.

Friedrich, M., & Friederici, A.D. (2006). Early N400 development and later language acquisition. *Psychophysiology, 43*, 1-12.

Friedrich, M., & Friederici, A.D. (2008). Neurophysiological correlates of online word learning in 14-month-old infants. *Neuroreport, 19*(18), 1757-1761.

Friedrich, M., & Friederici, A.D. (2010). Maturing brain mechanisms and developing behavioral language skills. *Brain and Language, 114*, 66-71.

Friedrich, M., & Friederici, A.D. (2011). Word learning in 6-month-olds: Fast encoding-weak retention. *Journal of Cognitive Neuroscience*, *23*(11), 3228-3240.

Friedrich, M., Wilhelm, I., Born, J., & Friederici, A.D. (2014). Generalization of word meanings during infant sleep. *Nature Communications, 6*, 1-9.

Fugati, Y., Ishihara, T., Tsude, K., Suzuki, Y., & Goto, M. (1998). Theta rhythms associated with sucking, crying, gazing and handling in infants. *Electroencephalography and Clinical Neuropsychology, 106*(5), 392-399.

Gallese, V. (2007). Before and below ‘‘theory of mind”: Embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society A,362*, 659-669.

Gallese, V. (2008). Mirror neurons and the social nature of language: the neural exploitation hypothesis. *Soc. Neurosci.* 3, 317–333.

Gallese, V., Fediga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition and the premotor cortex. *Brain, 119* (2*)*, 593-609.

Gampe, A., & Daum, M.M. (2014). Productive verbs facilitate action prediction in toddlers, *Infancy, 19*(3), 301-325.

Gampe, A., Brauer, J., & Daum, M.M. (2016). Imitation is beneficial for verb learning in toddlers. *European Journal of Developmental Psychology, 13*(5), 594-613.

Gentilucci, M., Benuzzi, F., Bertolani, L., Daprati, E., & Gangitano, M. (2000). Language and motor control. *Experimental Brain Research, 133*, 468-490.

Gergely, G., Bekkering, H., & Kiraly, I. (2002). Rational imitation in preverbal infants. *Nature*, 415, 755.

Gerson, S.A., Bekkering, H., & Hunnius, S. (2015). Short-term motor training, but not observational training alters neurocognitive mechanisms of action processing in infancy. *Journal of Cognitive Neuroscience, 27*(6), 1207-1214.

Glenberg, A. M., and Gallese, V. (2012). Action-based language: a theory of language acquisition, comprehension, and production. *Cortex* 48, 905–922.

Glenberg, A.M., & Kaschak, MP. (2002). Grounding language in action. *Psychological Bulletin & Review, 9 (3)*, 558-565.

Glenberg, A.M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *Quarterly Journal of Experimental Psychology, 61*, 905-919.

Glover, S., Rosenbaum, D.A. Graham, J., & Dixon, P. (2004). Grasping the meaning of words. *Experimental Brain Research, 154*, 103-108.

Golinkoff, R. M., Chung, H. L., Hirsh-Pasek, K., Liu, J., Bertenthal, B. I., Brand, R., et al. (2002). Young children can extend motion verbs to point-light displays. *Developmental Psychology, 38*, 604–615.

Gonzalez, S.L., Reeb-Sutherland, B.C., & Nelson, E.L. (2016). Quantifying motor experience in the infant brain: EEG power, coherence, and mu desynchronization. *Frontiers in Psychology, 7*, 216.

Graven, S.N., & Brown, J.V. (2008). Auditory development in the fetus and infant. *Newborn and Infant Nursing Reviews, 8*(4), 187-193.

Grezès, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation and verb generation of actions: a meta analysis. *Human Brain Mapping, 12*(1), 1-19.

Gunter, T.C., Stowe, L.A., & Mulder, G. (1997). When syntax meets semantics. *Psychophysiology, 34(6)*, 660-676.

Grossmann, T., Cross, E.S., Ticini, L.F., & Daum, M.M. (2012). Action observation in the infant brain: the role of body forma and motion. *Social Neuroscience, 8*(1), 22-30.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human and motor cortex. *Neuron, 41*(2), 301-307.

Henderson, L.M., Weighall, A.R., & Gaskell, M.G. (2013). Learning new vocabulary during childhood: Effects of semantic training on lexical consolidation and integration. *Journal of Experimental Child Psychology, 116*(3), 572-992.

Henderson, L.M., Weighall, A.R., Brown, H., & Gaskell, M.G. (2012). Consolidation of vocabulary is associated with sleep in children. *Developmental Science, 15*(5), 674-687.

Hoehl, S., & Wahl, S. (2012). Recording infant ERP data for cognitive research. *Developmental Neuropsychology, 37*(3), 187-209.

Hoehl, S., Michel, C., Reid, V.M., Parise, E., & Striano, T. (2014). Eye contact during live social interaction modulates infants' oscillatory brain activity. *Social Neuroscience, 9*(3*)*, 300-308.

Hofer, T., Hauf, P., & Aschersleben, G. (2005). Infant's perception of goal-directed actions performed by a mechanical device. *Infant Behavior and Development, 28*(4), 466-480.

Holle, H., & Gunter, T.C. (2007). The role of iconic gestures in speech disambiguation: ERP evidence. *Journal of Cognitive Neuroscience, 19*, 1175-1192.

Hommel, B. (2005). Perception and action: multiple roles of sensory information in action control. *Cognitive Processing, 6*(1), 3-14.

Hoorman, J., Falkenstein, M., Schwarzenau, P., & Hohnsbein, J. (1998). Methods of the quantification and statistical testing of ERP differences across conditions. *Behavioral Research Methods, Instruments, & Computers, 30*(1), 103-109.

Hunnius, S. & Bekkering, H. (2010). The early development of object knowledge: a study of infants' visual anticipations during action observation. *Developmental Psychology,* 46(2), 446-54.

Hunnius, S., & Bekkering, H. (2014). What are you doing? How active and observational experience shape infants action understanding. *Philosophical Transactions of the Royal Society: Biological Science, 369*(1644).

Iverson, J.M., & Goldin-Measow, S. (2005). Gesture paves the way for language development. *Psychological Science, 16*(5), 367-371.

Kaduk, K., Bakker, M., Juvrud, J., Gredebäck, G., Westermann, G., Lunn, J., & Reid, V.M. (2016). Semantic processing of actions at 9 months is linked to language proficiency at 9 and 18 months. *Journal of Experimental Child Psychology, 105*, 96-108.

Kaduk, K., Elsner, B., & Reid, V.M. (2013). Discrimination of animate and inanimate motion in 9-month-old infants: An ERP study. *Developmental Cognitive Neuroscience*, *6*, 14-22.

Kelly, A.M.C., & Garavan, H. (2005). Human functional neuroimaging of brain changes associated with practice. *Cerebral Cortex, 15*(8), 1089-1102.

Kelly, S.D., Ward, S., Creigh, P., & Bartolotti, J. (2007). An intentional stance modulates the integration of gesture and speech during comprehension. *Brain and Language, 101*, 222-233.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research Reviews, 29*(2-3), 169-195.

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, *16*(12), 606-617.

Klimesch, W., Doppelmayr, M., Pachinger, Th., & Ripper, B (1997). Brain oscillations and human memory: EEG correlates in the upper alpha and theta band. *Neuroscience Letters, 238*(1-2), 9-12.

Klimesch, W., Doppelmayr, M., Russegger, H., & Pachinger, T. (1996). Theta band power in the human scalp EEG and the encoding of new information. *Neuroreport, 17*(7), 1235-1240.

Klimesch, W., Pfurtscheller, G., Schimke, H. (1992). Pre- and post-stimulus processes in category judgement tasks as measured by event-related desynchronization (ERD). *Journal of Psychophysiology, 6*(3), 1992, 185-203.

Kochukhova, O. & Gredebäck, G. (2010). Preverbal infants anticipate that food will be brought to the mouth: an eye tracking study of manual feeding and flying spoons. *Child Development*. 81(6), 1729-38.

Kohler, E., Keysers, C., Umiltà, MA., Fogassi, L., Gallese, V., & Rozzolatti, G. (2002). Hearing sounds, understanding actions: action representations in mirror neurons. Science, 297 (5582), 846-848.

Kopp, F., & Lindenberger, U. (2011). Effects of joint attention on long-term memory in 9-month-old infants: an event-related potentials study. *Developmental Science, 14*(4), 660-672.

Kraljevic´, J., Cepanec, M., & Šimleša, S. (2014). Gestural development and its relation to a child's early vocabulary. *Infant Behavior and Development, 37*, 192-202.

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, 100*(15), 9096-9101.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual review of Psychology*, *62*, 621-647.

Kutas, M., & Hillyard, S.A. (1980). Event-related potentials to semantically inappropriate and surprisingly large words. *Biological Psychology, 11*, 99-116.

Kutas, M., & Hillyard, S.A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203-205.

Lachat, F., Hugueville, L., Lemarechal, J. D., Conty, L., & George, N. (2012). Oscillatory Brain Correlates of Live Joint Attention: A Dual-EEG Study. *Frontiers in Human Neuroscience*, 6, 156-165.

Legerstee, M. (1991). The role of person and object in eliciting early imitation. *Journal of Experimental Child Psychology*, *51*, 423-433.

Legerstee, M. (1992). A review of the animate-inanimate distinction in infancy: Implications for models of social and cognitive knowing. *Early Development and Parenting, 1(2)*, 59-67.

Leonard, C.H., & Hill, E.L. (2014). The impact of motor development on typical and atypical social cognition and language: a systematic review. *Child and Adolescent Mental Health, 19*(3), 163-170.

Libertus, K., & Violi, D. (2016). Sit to Talk: Relation between Motor Skills and Language Development in Infancy. *Frontiers in Psychology, 7*(475), 1-8.

Locatelli, M., Gatti, R., & Tettamanti, M. (2012). Training of manual actions improved language understanding of semantically related action sequences. *Frontiers in Psychology, 3*.

Ma., W., Golinkoff, R.M., Houston, D., & Hirsh-Pasek, K. (2011). Word learning in infant-and adult-directed speech. *Language Learning and Development, 7*, 209-225.

Mampe, B., Friederici, A.D., Christophe, A., & Wermke, K. (2009. Newborns' cry melody is shaped by their native language. *Current Biology, 19*(23), 1994-1997.

Mandler, J.M. (2006). Actions organize the infant's world. In K. Hirsh-Pasek & R. Michnick Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 111-133). New York: Oxford University Press.

Marshall, P.J., & Meltzoff, A.N. (2011). Neural mirroring systems: exploring the EEG mu rhythm in human infancy. *Developmental Cognitive Neuroscience, 1*(2), 110-123.

Marshall, P.J., Bar-Haim, Y., & Fox, NA. (2002). Development of the EEG from 5 months to 4 years of age. *Clinical Neurophysiology, 113* (8), 1199-1208.

Marshall, P.J., Young, T., & Meltzoff, A.N. (2011). Neural correlates of action observation and execution in 14-month-old infants: an event-related EEG desynchronization study. *Developmental Science, 14*(3), 474-480.

Martin, A., Wiggs, C.L., Ungerleider, L.G., & Haxby, J.V. (1996). Neural correlates of category-specific knowledge. *Nature, 379*, 649-652.

Marx, V., & Nagy, E. (2015). Fetal behavioural responses to maternal voice and touch. *PloS One, 10*(6), 1-15.

Mather, E., & Plunkett, K. (2012). The role of novelty in early word learning. *Cognitive Science, 36,* 1157-1177.

Mayer, M., Hunnius, S., van Elk., M., van Ede, F., & Bekkering, H. (2011). Joint action modulates motor system involvement during action observation in 3-year-olds. *Experimental Brain Research, 211*(3), 581-592.

Meltzoff, A.N. (1990). Foundations for developing a concept of self: the role of imitation in relating self to other and the value of social mirroring. In Cicchetti, Dante (Ed); Beeghly, Marjorie (Ed). (1990). *The self in transition: Infancy to childhood*, (pp. 139-164). Chicago, IL, US: University of Chicago Press, viii, 399 pp.

Meltzoff, A.N., & Moore, M.K. (1997). Explaining facial imitation: a theoretical model. *Early Development Parenting, 6*(3-4), 179-192.

Meltzoff, A.N., Murray, L., Simpson, E., Heimann, M., Nagy, E., Nadel, J., Pedersen, E.J., Brooks, R., Messinger, D.S., de Pascalis, L., Subiaul, F., Paukner, A., & Ferrari, P.F. (2017). Re-examination of Oostenbroek et al. (2016): evidence for neonatal imitation of tongue protrusion**,** *Developmental Science*, e12609.

Meltzoff. A.N., & Decety, J. (2003). What imitation tells us about social cognition: a reaproachment between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society, Biological Sciences, 358*(1431), 491-500.

Mesman, J., van IJzendoorn, M.H., & Bakermans-Kranenburg, M.J. (2009). The many faces of the Still-Face Paradigm: A review and meta-analysis. *Developmental Review, 29*(2), 120-162.

Michel, C., Stets, M., Parise, E., Reid, V.M., Striano, T., Hoehl, S. (2015). Theta- and alpha-band EEG activity in response to eye gaze cues in early infancy. *NeuroImage, 118*, 576-583.

Mills, D.L., Plunkett, K., Prat, C., & Schafer, G. (2005). Watching the infant brain learn words: Effects of vocabulary size and experience. *Cognitive Development, 20*(1), 19-31.

Mölle, M., Marshall, L., Fehm, H.L., & Born, J. (2002). EEG theta synchronization conjoined with alpha desynchronization indicate intentional encoding. *European Journal of Neuroscience, 15*(5), 923-928.

Moody, C.L., & Gennari, S.P. (2010). Effects of implied physical effort in sensorimotor and pre-frontal cortex during language comprehension. *NeuroImage, 49*, 782-793.

Moreno, I., De Vega, M., & Leon, I. (2013). Understanding action language modulates oscillatory mu and beta rhythms in the same way as observing actions. *Brain and Cognition, 82*(3), 236-242.

Morton, J., & Johnson, M.H. (1991) Conspec and conlern: a two process theory of infant face recognition. *Psychological Review, 98*(2), 164-181.

Myung, J., Blumstein, S.E., & Sedivy, J.C. (2006). Playing on the typewriter, typing on the piano: manipulation knowledge of objects. *Cognition, 98,* 223-242.

Ní Choisdealbha, Á., & Reid, V.M. (2014). The developmental cognitive neuroscience of action: semantics, motor resonance and social processing. *Experimental Brain Research*, *232*(6), 1585-1597.

Nishitani, N., & Hari, R. (2002). Viewing lip forms: cortical dynamics. *Neuron, 36,* 1211-1220.

Nordqvist, E., Rudner, M., Johansson, M., Lindgren, M., & Heimann, M. (2015). The relationship between deferred imitation, associative memory, and communication in 14-month-old children. Behavioral and electrophysiological indices. *Frontiers in Psychology, 6*(260).

Nyström, P., Ljunghammar, T., Rosander, K., & von Hofsten, C. (2010). Using mu rhythm desynchronization to measure mirror neuron activity in infants. *Developmental Science, 14*(2), 327-335.

Nyström, P., Ljunghammr, T., Rosander, K., & von Hofsten, C. (2008). Using mu rhythm desynchronization to measure mirror neuron activity in infants. *Developmental Science, 14*(2), 327-335.

Orekhova, E.V, Stroganova, T.A., & Posikera, I.N. (1999). Theta synchronization during sustained anticipatory attention in infants over the second half of the first year of life. *International Journal of Psychology, 32*(2), 151-172.

Orgs, G., Dombrowski, J.H., Heil, M., & Jansen-Osmann, P. (2008). Expertise in dance modulates alpha/beta event-related desynchronization during action observation. *European Journal of Neuroscience, 27*, 3380-3384.

Oostenbroek, J., Suddendorf, T., Nielsen, M., Redshaw, J., Kennedy-Constantini, S., Davis, J., & Slaughter, V. (2016). Comprehensive longitudinal study challenges the existence of neonatal imitation in humans, *Current Biology*, *26,* 1334-1338.

Oudgeneog-Paz. O., Volman, M.J.M., & Leseman, P.P.M. (2012). Attainment of sitting and walking predicts development of productive vocabulary between ages 16 and 28 months. *Infant Behavior and Development, 35*(4), 733-736.

Parise, E., & Csibra, G. (2012). Electrophysiological evidence for the understanding of maternal speech by 9-month-old infants. *Psychological Science*, *23*(7), 728-733.

Parise, E., Reid, V.M., Stets, M., Striano, T., (2008). Direct eye contact influences the neural processing of objects in 5-month-old infants. *Social Neuroscience* 3(2), 141-150.

Pascalis, O., de Haan, M., & Nelson, C.A. (2002). Is face processing specific during the first year of life. *Science, 296*(5571), 1321-1323.

Paulus, M., Hunnius, S., van Elk, M., & Bekkering, H. (2012). How learning to shake a rattle affects 8-month-old infants' perception of the rattle's sound:

Electrophysiological evidence for action-effect binding in infancy. *Developmental Cognitive Neuroscience, 2* (1), 90-96.

Piaget, J. (1952). *The origins of intelligence in children*. New York, NY: International Universities Press.

Pineda, J.A., (2005). The functional significance of mu rhythms: translating "seeing" and "hearing" into "doing". *Brain Research Reviews, 50*(1), 57-68.

Poulin-Dubois, D., Lepage, A. & Ferland, D. (1996). Infants' concept of animacy. *Cognitive Development, 11*, 19-36.

Prinz, W. (1990). A common coding approach to perception and action. In: Neumann O, Prinz W (eds) *Relationships between perception and action*. Springer, Berlin Heidelberg New York, pp 167–201.

Pulvermüller, F. (2005). Opinion: Brain mechanisms linking language and action. *Nature Reviews Neuroscience. 6,* 576-582.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Science, 22*(2), 253-279.

Pulvermüller, F., Härtle, M., & Hummel, F. (2001). Walking or talking? Behavioural and neurophysiological correlates of action verb processing. *Brain and Language, 78*(2), 143-168.

Rämä, P., Sirri, L., & Serres, J. (2013). Development of lexical-semantic language systems: N400 priming effect for spoken words in 18- and 24-month-old children. *Brain and Language, 125*, 1-10.

Reid, V. M., Hoehl, S., Grigutsch, M., Groendahl, A., Parise, E., & Striano, T. (2009). The neural correlates of infant and adult goal prediction: evidence for semantic processing systems. *Developmental Psychology*, *45*(3), 620-629.

Reid, V.M., & Striano, T. (2007). The directed attention model of infant social cognition. *European Journal of Developmental Psychology, 4*, 100-110.

Reid, V.M., & Striano, T. (2008). N400 involvement in the processing of action sequences. *Neuroscience Letters, 433*, 93-97.

Reid, V.M., Striano, T., Iacoboni, M. (2011). Neural correlates of dyadic interaction during infancy. *Developmental Cognitive Neuroscience, 1*(2), 124-130.

Reynolds, G.D., & Richards, J.E. (2005). Familiarization, attention, and recognition memory in infancy: An event-related potential and cortical source localization study. *Developmental Psychology, 41*, 598-615.

Richards, M. P. M. (1974). First steps in becoming social. In M. P. M. Richards (Ed.), *The Integration of a Child into a Social World*. Cambridge: Cambridge University Press.

Rizzolatti, G. (2005). The mirror neuron system and its function in humans. *Anatomy and Embryology, 210*(5), 419-421.

Rizzolatti, G. & Arbib, M.A. (1998). Language within our grasp. *Trends in Neuroscience, 21*(5), 188-194.

Rizzolatti, G., & Craighero, L. (2004). The mirror neuron system. *Annual Review of Neuroscience, 27*, 169-192.

Rizzolatti, G., Ferida, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3* (2), 131-141.

Rowe, M.L., & Goldwin-Meadow, S. (2009a). Differences in early gesture explains SES disparities in child vocabulary size at school entry. *Science, 323*, 951-953.

Rowe, M.L., & Goldwin-Meadow, S. (2009b). Early gesture selectively predicts later language learning. *Developmental Science, 12*, 182-187.

Rüschemeyer, S.A., Lindemann, O., van Elk, M., & Bekkering, H. (2009). Embodied cognition: the interplay between automatic resonance and selection-for-action mechanisms. *European Journal of Social Psychology, 39*(7), 1180-1187.

Rüschemeyer, S.A., Brass, M., & Friederici, A.D. (2007). Comprehending prehending: neural correlates of processing verbs with motor stems. *Journal of Cognitive Neuroscience, 19*(5), 855-865.

Saby, J.N., Marshall, P.J., & Meltzoff, A.N. (2012). Neural correlates of being imitated: an EEG study in preverbal infants. *Social Neuroscience, 7*(6), 650-661.

Sacco, K., Cauda, F., Cerliani, L., Mate, D., Duca, S., & Geminiani, G.C. (2006). Motor imagery of walking following training in locomotor attention. The effect of "the tango lesson". *NeuroImage, 32(3),* 1441-1449.

Sai, F.Z. (2004). The role of the mother's voice in developing mother's face preference: Evidence for intermodal perception at birth. *Infant and Child Development, 14*(1), 29-50.

Senju, A., & Johnson, M.H. (2009). The eye contact effect: mechanisms and development. *Trends in Cognitive Sciences, 13*(3), 127-134.

Senju, A., Csibra, G., & Johnson, M.H. (2008). Understanding the referential nature of looking: infants' preference for object-directed gaze. *Cognition, 108*(2), 303-319.

Shukla, M., White, K.S., & Aslin, R.N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Science*, *108*(15), 6038–6043.

Singh, L., Nestor, S., Parikh, C., & Yull, A. (2009). Influences of infant-directed speech on early word recognition. *Infancy, 14*(6), 654-666.

Snyder, K., Webb, S.J., & Nelson, C.A. (2002). Theoretical and methodological implications of variability in infant brain response during a recognition memory paradigm. *Infant Behavior and Development, 25*, 466-494.

Sodestrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review, 27*(4), 501-532.

Sommerville, J. & Woodward, A.L. (2005). Pulling out the intentional structure of action: the relation between action processing and action production in infancy. *Cognition*, 95(1), 1-30.

Soska, K. C., Adolph, K. E., Johnson, & Scott P. (2010). Systems in development: Motor skill acquisition facilitates three-dimensional object completion. *Developmental Psychology, 46*(1), 129-138

Stets, M., & Reid, V.M. (2011). Infant ERP amplitudes change over the course of an experimental session: implications for cognitive processes and methodology. *Brain and Development*, *33*(7), 558-568.

Stets, M., Stahl, D., & Reid, V.M. (2012). A meta-analysis investigating factors underlying attrition rates in infant ERP studies. *Developmental Neuropsychology, 37*, 226-252.

Striano, T., Chen, X., Cleveland, A., & Bradshaw, S. (2006). Joint attention social cues influence infant learning. *European Journal of Developmental Psychology, 3*(3), 289-299.

Striano, T., Reid, V.M., & Hoehl, S. (2006). Neural mechanisms of joint attention in infancy. *European Journal of Neuroscience, 23*(10), 2819-2823.

Stroganova T.A., & Orekhova, E.V. (2007). EEG and infant states. In M. de Haan (Ed.), *Infant EEG and Event-Related Potentials* (pp. 251–287). Psychology Press: Hove and New York.

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology, 32*, 492–504.

Thomas, M.S.C., Annaz, D., Ansari, D., Serif, G., Jarrold, C., & Karmiloff-Smith, (2009). Using developmental trajectories to understand developmental disorders. *Journal of Speech, Language, and Hearing Research, 52*, 336-358.

Tomasello, M. (1995). In: Joint attention: Its origin and role in development, ed. C. Moore & P. J. Dunham, pp. 103–30. Erlbaum.

Tomasello, M., & Farrar, M.J. (1986). Joint Attention and Early Language. *Child Development, 57*(6), 1454-1463.

Tomasello, M., & Kruger, A. C. (1992). Acquiring verbs in ostensive and nonostensive contexts. *Journal of Child Language, 19*, 311–333

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development, 78*, 705-722.

Torkildsen, J., Friis Hansen, H., Scangstu, J.M., Smith, L., Simonsen, H.G., Moen, I., et al. (2009). Brain dynamics of word familiarization in 20-month-olds: Effects of productive vocabulary size. *Brain and Language, 108*, 73-88.

Torkildsen, J., Sannerud, T., Syversen, G., Thormodsen, R., Simonsen, H.G., Moen, I., Lindgren, M. (2006). Semantic organization of basic-level words in 20-month-olds: An ERP study. *Journal of Neurolinguistics, 19*, 431-454.

Twomey, K. E., Ranson, S. L., & Horst, J. S. (2014). That's more like it: Multiple exemplars facilitate word learning. *Infant and Child Development*, *23*(2), 105–122.

Uithol, S., van Rooji, R., Bekkering, H., & Haselager, P. 2(011). Understanding motor resonance. *Social Neuroscience, 6*(4), 388-397.

Urgesi, C., Moro, V., Candidi, M., & Aglioti, M.S. (2006). Mapping implied body actions in the human motor system. *The Journal of Neuroscience, 26* (30), 7942-7949.

van Elk, M., van Schie, H.T., Hunnius, H., Vesper, C. & Bekkering, H. (2008). You'll never crawl alone: Neurophysiological evidence for motor resonance in infancy. *Neuroimage, 43(4),* 808-814.

van Elk, M., van Schie, H.T., Zwaan, R.A., Bekkering, H. (2010). The functional role of motor activation on language processing: motor cortical oscillations support lexical-semantic retrieval. *NeuroImage, 50*, 665-677.

Voght, S., Buccino, G., Wohlschlager, A.M., Canessa, N., Shah, N.J., Zilles, K., Eickhoff, S.B., Freund, H-J., Rizzolatti, G., & Fink, G.R. (2007). Prefrontal involvement in imitation learning of hand actions: effect of practice and expertise. *NeuroImage, 37*(4). 1371-1383.

von Hofsten, C. (2004). An action perspective on motor development. Trends in *Cognitive Sciences, 8*(6), 266-272.

Walle, E.A., & Campos, J. (2014). Infant language development is related to the acquisition of walking. *Developmental Psychology, 50*(2), 336-348.

Ward, LM. (2003). Synchronous neural oscillations and cognitive processes. *Trends in Cognitive Sciences, 7*(12), 553-559.

Webb, S.J., Long, J.D., & Nelson, C.A. (2005). A longitudinal investigation of visual event-related potentials in the first year of life. *Developmental Science, 8*, 605-616.

Wei, G., & Luo, J. (2010). Sport expert's motor imagery: functional imaging of professional motor skills and simple motor skills. *Brain Research, 1341*, 52-62.

Wojcik, E.H. (2013). Remembering new words: integrating early memory development into word learning. *Frontiers in Psychology, 4(151)*, 1-12.

Wu, Y.C., & Coulson, S. (2005). Meaningful gestures: Electrophysiological indices of iconic gesture comprehension. *Psychophysiology, 42*, 654-667.

Zacks, J.M., Braver, T.S., Sheridan, M.A., Donaldson, D.I., Snyder, A.Z., Ollinger, J. M., Raichle, M.E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience, 4*, 651-655.

Zangl, R., & Mills, D.L., (2007). Increased brain activity to infant-directed speech in 6- and 13-month-old infants. *Infancy, 11*(1), 31-62.

Zoia, S., Blason, L., D'Ottavio, G., Bulgheroni, M., Pezetta, Aldo, S., & Casteillo, U. (2007). Evidence of early development of action planning in the human foetus: a kinematic study. *Experimental Brain Research, 176*(2), 217-226.

Zwaan, R.A. & Taylor, L.J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *Journal of Experimental Psychology: General, 135*, 1-11.